APPENDIX E-2
Record Content Analysis
Findings, Discussion, and Recommendations

APPENDIX E-2
Record Content Analysis
Findings, Discussion, and Recommendations

Table of Contents

APPENDIX E-2
Record Content Analysis
Findings, Discussion, and Recommendations

1.0.   INTRODUCTION

Content analysis of GILS records served three purposes: to assess records' quality in terms of completeness and accuracy; to explore the relationship of selected characteristics of records and serviceability in networked information discovery and retrieval (NIDR); and to develop recommendations for future application or adaptation of the method.

More than 3500 instances of metadata were evaluated for incidence and/or content, and entered into a database for subsequent coding and analysis. This appendix presents the results of that analysis, along with a discussion of implications and recommendations. In addition, the investigators maintained a log of areas for further research (see Section 8.0) that may be utilized by system developers, specification and procedures writers, and those with direct responsibility for GILS record quality.

1.1.   Organization of Material

Section 3.0 Data Summaries aggregates significant results in terms of the analysis objectives. Detailed results of the analysis, Section 4.0 Findings, Discussion, and Recommendations, are presented in four major categories, which were assigned alphabetic codes. "Accuracy (A)" concerns the incidence of errors. "Completeness (C)" includes data concerning GILS record element utilization and values. "Resource Profile (P)" includes findings concerning general characteristics, such as aggregation and objects represented, for the sample population. The final section, "Serviceability (S)," presents findings relevant to record effectiveness in NIDR and user convenience. Further, each category code was coupled with a numeric code that reflects the order of data collection; under "Completeness" the data are additionally sorted in preferred element display order [per *Federal information processing standards publication 192, Application Profile for the Government Information Locator Service (GILS)* (National Institute for Standards and Technology, 1994) and *The government information locator service: Guidelines for the Preparation of GILS Core Entries* (National Archives and Record Administration, 1995a)] for ease of reference.

The following table offers a summary of the organizational scheme used in Section 4.0 Findings, Discussion, and Recommendations.

Table E2-1
Organization of Results

2.0.   METHOD OVERVIEW

The analysis was performed in two phases: Phase 1 comprised examination of a pool of 80 records from 40 agencies' GILS retrieved deliberately to represent a range of information resource types (e.g., databases, catalogs, records systems). These records served as the basis for developing and operationalizing a set of more than 50 qualitative and quantitative evaluative criteria that included records' format, aggregation, media representation, and descriptiveness. Descriptiveness was defined as the incidence of utilization and content (value) attributes for all mandatory and selected optional elements and subelements specified by *FIPS Pub. 192* Annex E-GILS Core Elements definitions as reproduced and supplemented by usage guidelines and examples in the NARA *Guidelines*. In Phase 2, these criteria were systematically applied to a set of 83 records randomly retrieved January 13 and 14, 1997, from 42 agencies' GILS. Results, therefore, reflect record content *at the time of retrieval* and represent a "snapshot" during only one, and arbitrary, point in the GILS system lifecycle. In addition, the "Core subset" analysis comprises records within the total sample that contained a value of "US Federal GILS" or "U.S. Federal GILS" in the Controlled Vocabulary-Local Subject Index Term subelement.

Appendix C-4 Record Content Analysis Methodology provides a complete discussion of methodology, and Appendix D-4 Record Content Analysis Instrument presents the database fields used for data collection.

3.0.   DATA SUMMARIES

The following tables summarize significant data resulting from the analysis in terms of the objective-based categories outlined in Section 1.1 Organization of Material. Operational definitions of semantics used during the analysis are presented in Appendix C-4 Record Content Analysis Methodology and reiterated as applicable in Section 4.0 Findings, Discussion, and Recommendations. Investigators strongly recommend that interpretation of the following findings be guided by the complete data and discussions provided in Section 4.0.

Section 3.5 High Quality Records From the Sample cites records exemplifying many of the important quality characteristics. These records are reproduced (as printed directly from the Web) in Attachments E2-1a through E2-1d to this appendix.

3.1.   Accuracy

The following summary data reflect the number of errors identified in the sampled GILS records. (The scope of the current study did not include verification of the "accuracy" of substantive information *per se*—such as telephone numbers, URLs, etc.)

Criteria	Data Summary and Highlights	Section 4.0
Records with incorrect date formats	33%	C3.16
Records with file formatting errors	24%	A1
Records with spelling or typographical errors	10%	A2

3.2 Completeness

This table presents summary data concerning the fullness of sampled records in terms of inclusion of both "mandatory" and "optional" elements (as defined by the NARA *Guidelines*).

Criteria	Data Summary and Highlights	Section 4.0
General		
Number of populated elements per record	max 190 min 11 avg 42	C1
Records containing "blank" (labeled but null value) elements	36%	C2
Locally defined elements identified	12	S8
Utilization of 12 mandatory elements (excludes AIS-Mandatory and RM-dependent)		
Total sample	96%	C3.1-16
Core subset	95%	C3.1-16
Utilization of selected optional elements		
Controlled Vocabulary	12% LCSH most popular thesaurus	C4.1 C4.2
Local Subject Index	54% (includes variants of "US Federal GILS")	C4.3
Availability-Resource Description	12% for total sample 24% for core subset	C4.4
Methodology	2%	C4.5
Originator subelement(s)	65%	C3.2
Record Source subelement(s)	63%	C3.15

3.3. Resource Profile

The following table summarizes characteristics of the resources described in the sampled records.

Criteria	Data Summary and Highlights	Section 4.0
Record types (AIS, Locator, Privacy Act systems)	could not be discerned	P1
Objects represented	(12 types) 22% "subject matter database" 19% "publication" 59% other	P2
Aggregation	(5 levels) 36% Record aggregated objects 25% Aggregated object represented 20% Discrete object 12% Object aggregates metadata 6% Unknown	P3
Containers	(7 types) 22% Multiple 23% Print 8% Web	P4

3.4.   Serviceability

The "serviceability" data summarized below are considered to represent record effectiveness in terms of the degree to which they enhance NIDR, convenience to the user, aesthetics, readability, and relevance judgment.

Criteria	Data Summary and Highlights	Section 4.0
NIDR Factors		
Records with spelling or typographical errors	10%	A2
Records with Controlled Vocabulary	12%	C4.1
	LCSH most popular thesaurus	C4.2
Records with Local Subject Index	54%	C4.3
	6% of Core subset contained "U.S." rather than "US" in <US Federal GILS>	C3.3
Records with Resource Description	12% for total sample	C4.4
	24% for core subset	
Records with Record Schedule number	14%	C3.13
Records with Control Identifier	91%	C3.14
Records with Originator subelement(s)	65%	C3.2
Records with Record Source subelement(s)	63%	C3.15
Locally defined elements identified	12	S9
Aggregation	36% Record aggregated objects	P3
	25% Aggregated object represented	
	20% Discrete object	
	12% Object aggregates metadata	
	6% Unknown	
Records with (any) acronyms	65%	S6
Records with undefined acronyms	12%	S6
Records with dates in incorrect format	33%	C3.16
Records with legislative citation	48% (only one was GILS-related)	S7
User Convenience		
Preferred element display order	64% for total sample	S5
	57% for Core subset	
Points of Contact	50% offices	C3.12
	23% personal names	
	9% job title	
	3% other	
Records with Availability-Distributor	93%	C3.7
Availability-Order Process	86%	C3.8
Records with hypertext	25% total	S2
	52% in Available Linkage	
	24% in Distributor Network Address	
File formats	81% ASCII	S1
	83% HTML	
	2% SGML	
	0% PDF	

(continued)

Criteria	Data Summary and Highlights	Section 4.0
Aesthetics/Readability		
Number of populated elements per record	max 190 min 11 avg 42	C1
Records containing "blank" (labeled but null value) elements	36%	C2
File formats	81% ASCII 83% HTML 2% SGML 0% PDF	S1
Records with file formatting errors	24%	A1
Capitalization style	86% sentence-case 10% elements-only capitalized	S3
Records employing any indentation pattern	73%	S4
Relevance-Judgment Factors		
Records with descriptive Titles	75%	C3.1
Records with descriptive Abstract	86%	C3.4
Records with (any) undefined acronyms	12%	S6
Records with substantive Access Constraints	29%	C3.10
Records with substantive Use Constraints	17%	C3.11
Records naming container (dissemination media)	50%	P4
Locally defined elements identified	12	S8
Records with Resource Description	12% for total sample 24% for core subset	C4.4
Records misusing Date of Last Modification as referring to resource rather than record	at least 4	C3.16

3.5. Examples of High-Quality Records from Sample

Four records exhibiting characteristics of "high quality" are provided as examples in Attachments E2-1a through E2-1d to this appendix.

- *AHCPR Publications Clearinghouse* available at <http://www.dhhs.gov/progorg/oirm/newhhsgils.htm> by searching Control Identifier (quotes required) "HHS-AHC-00509"
- *Aviation Accident Synopses World Wide Web Page* available by browsing <http://www.ntsb.gov/Info/Info.htm> or directly at <http://www.ntsb.gov/Info/GILS/GILSSYN.htm>
- *Farm Credit Administration's Privacy Act Systems* available <http://www.access.gpo.gov/su_docs/gils/gilsfld.html> by searching Control Identifier (quotes required) "FCA/PA-1"
- *FEMA Publications Catalog* available <http://www.access.gpo.gov/su_docs/gils/gils.html> by searching Federal Emergency Management Agency for "FEMA0001"

These records contain mandatory elements populated with NARA *Guidelines*-compliant values and are highly readable and descriptive without excessive length. In addition, they represent a range of "information objects" and "containers" (see Appendix C-4 Record Content Analysis Methodology): an information resource organization, an aggregated set of reports available via Web site, a "system of records" available via Government Printing Office (GPO) GPO Access, and a traditional printed publications catalog, respectively.

4.0.   FINDINGS, DISCUSSION, AND RECOMMENDATIONS

Detailed results of the analysis are presented below in four major categories, which were assigned alphabetic codes. Section 4.1 "Accuracy (A)" concerns the incidence of errors. Section 4.2 "Completeness (C)" includes data concerning GILS record element utilization and values. Section 4.3 "Resource Profile (P)" includes general characteristics of the records studied. The final section, Section 4.4 "Serviceability (S)," presents findings relevant to record effectiveness in NIDR and user convenience. Further, each category code was coupled with a numeric code that reflects the order of data collection; under "Completeness" the data are additionally sorted in preferred element display order (per *FIPS Pub. 192* and the NARA *Guidelines*) for ease of reference.

4.1.   Accuracy

The following data reflect the number of errors identified in the sampled GILS records. (The scope of the current study did not include verification of the "accuracy" of substantive information *per se*—such as telephone numbers, URLs, etc.) File formatting errors were found in about 25% of sampled records; spelling and typographical errors were evident in 10%.

A1   Formatting Errors
Findings: Roughly 1 in 4 records sampled contained error(s) attributable to file formatting or conversion, such as no hard-returns, unintentional hard-returns, incomplete files, HTML tags, stray ASCII text, file format characters, etc.

TOTAL SAMPLE

FILE FORMAT ERRORS	N	%
YES	20	24%
NO	63	76%
TOTAL	83	100%

CORE SUBSET

FILE FORMAT ERRORS	N	%
YES	7	17%
NO	35	83%
TOTAL	42	100%

Discussion: These data support record-creator complaints during focus groups and interviews with key informants that support personnel at times/places do not have online access with which to view GILS product. Study participants (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations) showed poor tolerance of formatting errors.
Recommendations: Devise a hard-/software independent template and/or HTML editor for record formatting, or limit formatting responsibility to agency or subcontracted personnel with Web browsers.

A2   Spelling and Typographical Errors
Findings: The examination revealed that 1 in 10 records sampled contained spelling or typographical errors; the percentage was somewhat lower in the Core subset.

TOTAL SAMPLE

SPELLING ERRORS/TYPOS	N	%
YES	8	10%
NO	75	90%
TOTAL	83	100%

CORE SUBSET

SPELLING ERRORS/TYPOS	N	%
YES	3	7%
NO	39	93%
TOTAL	42	100%

Discussion: In addition to the possibility of record retrieval failure caused by lack of exact matching of user input, the presence of spelling and typographical errors may erode users' trust in the transmission and/or content integrity of records.
Recommendations: Use machine-based spell checkers, or assign checking responsibility to someone other than the writer.

4.2.   Completeness

The following data concern the fullness of sampled records in terms of inclusion of both "mandatory" and "optional" elements (as defined by the NARA *Guidelines*). Of 67 possible element tags, the sample averaged 42 elements containing substantive values of a value of "none," "not applicable," etc. Nearly 40% of the records sampled featured the practice of presenting some "blank" elements—i.e., labels with no data.

C1   Number of Elements Per Record

Findings: In the total sample, records averaged a "length" of 42 elements, with a maximum of 190 and minimum of 11. The most frequent element count (mode) was 33, featured by 8% of the records. The Core subset's element utilization rates were not significantly different from the total sample.

TOTAL SAMPLE

TOTAL FIELDS USED		
SUM	3500	
AVG	42	
MAX	190	
MIN	11	
MODE	33	8%

CORE SUBSET

TOTAL FIELDS USED		
SUM	1873	
AVG	45	
MAX	190	
MIN	11	
MODE	33	10%

Discussion: *FIPS Pub. 192* currently specifies 67 tags. User perception of a record's "length" may be related to several factors, including C2-Practice of Presenting Blank Elements, S1-File Formats, S3-Capitalization, S4-Indentation, and S5-Element Display Order, as well as the number of elements utilized and repeated and the extent of their values. In the Scripted Online User Assessment, a record containing 14 (total) populated elements was judged "just right" in length by most users; one user felt that record space was generally "wasteful in relation to what you get" clarified as meaning "not actual documents." Of the 190 elements of the maximum-length record noted above, 170 were Point of Contact subelements in which 17 agency field stations were described (for a "records management program"). In addition, although this practice was not assessed systematically, the investigators noted several instances of pronounced redundancy in record content—e.g., one record contained "browser is required" or equivalent in four different elements.

Recommendations: Implement Z39.50-compliant clients which will enable the presentation of customized views of the record. To address the appropriateness of content and placement of values, a possible research project might isolate a random sample of selected data values and test whether GILS-cognizant vs. noncognizant users can place them in the "correct" (per published standards) elements. For those not trained in bibliographic control, and even those who are familiar with library cataloging processes, the NARA *Guidelines* are complex and at times ambiguous. Record creators as well as quality checkers may require a more straightforward, procedural, or specific instructions as well as a FAQ list, pocket guide reference, context-sensitive online help, etc. to select elements required to describe the resource and its availability appropriately.

C2   Practice of Presenting "Blank" Elements

Findings: More than one-third of the records examined presented one or more elements (labels) containing no data. No pattern was discernible across agencies as to the basis for this practice. However, it is possible that record creators are working with a "generic," inflexible in the interest of time economy or internal quality assurance initiatives.

TOTAL SAMPLE

"BLANK" ELEMENTS	N	%
YES	30	36%
NO	53	64%
TOTAL	83	100%

CORE SUBSET

"BLANK" ELEMENTS	N	%
YES	17	40%
NO	25	60%
TOTAL	42	100%

Discussion: See C1-Number of Elements Per Record for a note about user perceptions of record "length." In addition, users may perceive "blank" elements as agency negligence or system error.

Recommendation: A further analysis of how record creators are handling elements perceived as irrelevant or not necessary (i.e., the incidence of "none," "not applicable," "N/A," "not required" and similar null values vs. presentation of "blank" elements vs. omission of such elements altogether) may indicate that a procedural standard is in order. A useful adjunct to this research could assess the incidence of elements presented for "fielded searching" and agency rationale for selection.

C3   Utilization and Selected Characteristics of "Mandatory" Elements

Overall, utilization of GILS mandatory elements was very high. Excluding Agency Program, Sources of Data, and Schedule Number because the incidence of AIS as a record type could not be determined, the analysis revealed 96% utilization for the total sample and 96% for the Core subset. Given that 50% of the records featured some variant of "US Federal GILS" in the Local Subject Index, this close match may be interpreted to mean either a nearly ubiquitous appreciation of the intrinsic value of the mandatory elements or a similarly ubiquitous uncertainty as to the designation "core" record.

The characteristics of *values* found in mandatory elements was less positive, however:

- The incidence of descriptive titles was relatively low—only 75% for the total sample and 67% for the Core
- Descriptiveness of Abstracts was slightly better than that of Titles, at around 86%
- Substantive Access Constraints and Use Constraints were named in 29% and 17% of the records, respectively
- Almost 25% of records sampled named an individual as Point of Contact
- The granularity of organizational descriptions varied widely; Originator and Record Source subelement(s) were used in about 65% of records
- The analysis revealed that nearly 40% of the records contained at least one date in other than YYYMMDD format.

These findings indicate that adherence to the NARA *Guidelines* in terms of element inclusion is high but that application of usage recommendations is less predictable.

The following results and discussion are presented per element in "preferred display" order for a GILS record (*FIPS Pub. 192* and NARA *Guidelines*). Note that the term utilization here means that the element was present, presented, and populated (even in cases where "none", "n/a", or similar acknowledgment of an absence of substantive content was present).

C3.1 Title

Findings: All records sampled utilized Title. Of the total sample, 3 in 4 records' Titles were coded descriptive; the incidence was slightly lower (67%) in the Core subset. The code of nondescriptive was applied to instances such as:

- Annual Reports failing to name the year
- a record titled "Employment Center," where the information object described was a telephonic job line
- use of the singular (e.g., "Report") when other element values implied more than one information object
- records titled simply "General Files" and "Minutes."

Examples of descriptive titles include "Investment Funds Brochure," "Automated Tariff Filing and Information System," and "GPO Access User Guide Online via GPO Access." In addition, although time constraints precluded a systematic assessment, the investigators noted at least the following terms used in titles of records describing an agency Web site: "Home Page," "Homepage," "World Wide Web Site," and "Internet Site."

TOTAL SAMPLE

TITLE DESCRIPTIVE	N	%
YES	62	75%
NO	21	25%
TOTAL	83	100%

CORE SUBSET

TITLE DESCRIPTIVE	N	%
YES	28	67%
NO	14	33%
TOTAL	42	100%

Discussion: The importance of descriptive Titles cannot be overemphasized given that it represents the only substantive content returned by most GILS search engines in the results list. The Scripted Online User Assessment determined that most users will judge a "hit's" relevancy first by appearance of their search terms in the Title, followed by WAIS "score," a finding related to a difficulty encountered during the current content analysis—particularly of titles of automated information systems. The NARA Guidelines state:

> [1]This element conveys the most significant aspects of the referenced resource and is intended for initial presentation to users independently of other elements. It should provide sufficient information to allow users to make an initial decision on likely relevance. It should convey the most significant information available, including the general topic area, as well as a specific reference to the subject.
> [2]The title provides the name of the information resource as assigned by the agency. For automated information systems and locators to information dissemination products, the title is the officially assigned name for the system.

This guideline assumes that the names of automated information systems (see P5-Object Represented) are descriptive for purposes of NIDR—an assumption the investigators find unwarranted in experience (the choice of name for a system is often motivated by its ability to create a facile acronym) and by way of this investigation. For example, are "OpenNet," "Enterprise Information System: EIS," "OEPC BBS," and "HUD USER," *descriptive* titles? According to [1] above, no. NARA, in their record "CLIO, The National Archives Information Server" attempts to accommodate both principles (and the investigators assume, although the term is not defined in the record, that "CLIO" spelled out might not be "descriptive."). The latter example brings to light another issue—that of search engines failure to return the agency acronym with the search result (title). This further degrades the NIDR value of titles such as "Consumer Bulletins," which emphasize audience at the expense of "the general topic area, as well as a specific reference to the subject" (per [2] above)—a problem that appears to be exacerbated upon wide cross-agency searching. Unfortunately, on the other side of this coin are titles such as "Federal Communications Commission (FCC) Technical Documents Created By Its Office of Engineering and Technology (OET)," a title that some might characterize as verbose. Section S2-Content Hypertext also discusses Title values.

Recommendations: Given that the title is the only record-content cue provided to current GILS users, it is essential that some degree of objective, third-party (i.e., other than resource creator and/or record creator) evaluation be applied when evaluating title descriptiveness. Implementation of Z39.50-compliant systems could eliminate this problem by allowing the user to select presentation of the Abstract with the Title to assist in judging relevancy.

C3.2 Originator

Findings: All records examined utilized the Originator element. 65% of the records sampled utilized at least one Originator subelement but only 45% presented its label.

TOTAL SAMPLE

ORIGINATOR SUBELEMENTS	N	%
YES	54	65%
NO	29	35%
TOTAL	83	100%

CORE SUBSET

ORIGINATOR SUBELEMENTS	N	%
YES	24	57%
NO	18	43%
TOTAL	42	100%

TOTAL SAMPLE

ORIGINATOR SUBELEMENTS LABELED	N	%
YES	37	45%
NO	46	55%
TOTAL	83	100%

CORE SUBSET

ORIGINATOR SUBELEMENTS LABELED	N	%
YES	16	38%
NO	26	62%
TOTAL	42	100%

Discussion: The NARA *Guidelines* provide the following definition of this element: "This element occurs once per locator record. It identifies the information resource originator, named as in the *U.S. Government Manual* where applicable." Most study participants felt that "all GILS records should look alike" (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations). It may be concluded that this preference refers to the presence and absence of display characteristics rather than content. An assessment of use of agency names as set forth in the *U.S. Government Manual* was not performed due to time constraints, however it is noted that this requirement will not serve the stated purpose of supporting NIDR unless the user has ready access to the *Government Manual* Table of Contents or Appendix A: Commonly Used Abbreviations and Acronyms.

Recommendations: The term "information resource originator" is undefined and its relationship to Point of Contact's "organizational unit that *created and maintains* [investigators' emphasis] the information dissemination product or information system" is unclear. It is suggested that research be conducted to assess users' expectations concerning presentation of true "authorship" data in the bibliographic sense as opposed to the entity responsible for compilation, administrative maintenance, or dissemination of the resource. In addition, it is suggested that a cross-tabulation of Originator/Point-of-Contact/Record Source values be performed and the values sampled for accuracy to verify potential confusion among definitional terms and roles implied by "originator," "creator," "provider," etc. (See discussions at C3.12-Point of Contact and C.15-Record Source.) Further assessment of user preferences for display of organizational subelement labels and values, as well as their aggregation levels and resultant maintenance burden, is recommended as well. Investigators believe that standardization of element label display will contribute to users' "footing" within GILS vs. other information space—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. In addition, on GPO Access GILS, it may be prudent to provide a hypertext link from "*US Government Manual*" in the field definition files to a recast version of *Government Manual* Appendix A: Commonly Used Abbreviations and Acronyms.

C3.3 Local Subject Index—"US Federal GILS"

Findings: As noted at the beginning of this section, for purposes of the GILS Evaluation record-content analysis, records containing "US Federal GILS" or "U.S. Federal GILS" in this element constituted the analysis subset called "Core subset." The NARA-recommended "US" format appeared in 43% of the records sampled.

TOTAL SAMPLE

US FED GILS	N	%
US FEDERAL GILS	36	43%
U.S. FEDERAL GILS	5	6%
NOT USED	41	49%
"US" BUT IN WRONG ELEMENT	1	1%
TOTAL	83	100%

CORE SUBSET

US FED GILS	N	%
US FEDERAL GILS	36	86%
U.S. FEDERAL GILS	5	12%
NOT USED	0	0%
"US" BUT IN WRONG ELEMENT	1	2%
TOTAL	42	100%

Recommendations: Investigators recommend that the concept and functionality of the GILS "Core" be re-examined in light of the study's overall findings and given the lack of significant difference in utilization and quality of "mandatory" vs. "optional" elements revealed during this content analysis. Should the requirement remain viable, the results above call for a clarification of purpose and implementation guidelines in GILS standards and procedures.

C3.4 Abstract

Findings: Only 2 records in the sample failed to provide a value in this element, one of these was in the Core subset. More than 85% of sampled records Abstracts were coded "descriptive." "Nondescriptive" incidents included:
- "none."
- "The [agency] is responsible."
- Values equaling the name or title of the information resource
- Investigator judgment that content matched more closely another element's definition (e.g., Purpose or Availability)
- Investigator judgment that the content was degraded by use of incomplete sentences or technical jargon or "bureaucrat-ese" (e.g., "It discusses the application of one or more provisions of law to the detailed factual situation set forth as a proposed course of conduct in light of requirements of a particular…regulation rule [sic], order to cease and desist or affirmative order…").

TOTAL SAMPLE

ABSTRACT DESCRIPTIVE	N	%
YES	71	86%
NO	10	12%
NOT USED	2	2%
TOTAL	83	100%

CORE SUBSET

ABSTRACT DESCRIPTIVE	N	%
YES	37	88%
NO	4	10%
NOT USED	1	2%
TOTAL	42	100%

Discussion: Per the NARA *Guidelines*:

> This [element's] narrative should provide enough general information to allow the user to determine if the information resource has sufficient potential to warrant contacting the provider for further information…The content of the abstract will be dependent upon the nature of the entity to be described (i.e., a locator to information dissemination products, a Privacy Act system, or an automated information system). The abstract may include, but is not limited to, discussion of the information content (including data coverage, persons, events, and topics); form of information; media; time span; and geographic coverage.

The Abstract is a familiar and preferred basis of relevancy judgment for many users. Creation of descriptive abstracts is time-consuming and especially difficult for personnel unfamiliar with the resource subject and/or abstracting principles.

Recommendations: More specific guidelines, and perhaps exposure to a greater number of effective (model) Abstracts for various information objects (see P5-Information Object), may assist record creators in developing consistency in the descriptions. In addition, investigators recommend that the definition, usage, and structure of the Resource Description subelement be revised to provide physical description of the object as recognizable by the *user*

rather than by the distributor, that this information be removed from the Abstract element to the Resource Description subelement, and that the subelement be mandatory and structurally associated with the Abstract element. See C4.4 Availability-Resource Description.

C3.5 Purpose
Findings: Of the total sample, 96% of records utilized this element; of the Core subset, 95% did.

TOTAL SAMPLE

PURPOSE	N	%
YES	80	96%
NO	3	4%
TOTAL	83	100%

CORE SUBSET

PURPOSE	N	%
YES	40	95%
NO	2	5%
TOTAL	42	100%

Discussion: The NARA *Guidelines* state that this element "describes why the information   resource is offered and identifies other programs, projects, and legislative actions wholly or partially responsible for the establishment or continued delivery of this information resource." S7-Citation of Legislation presents data concerning the total (element-wide) incidence; the NARA Guidelines also call for legislative references in the Agency Program element (see C3.6-Agency Program).

Recommendations: The feasibility of automating the insertion of substantively correct and properly formatted legislative citations upon inclusion of a program or project name (i.e., installation of an expert-system legislation index cross-reference macro) could reduce record creation and maintenance burden as well as facilitate identification of legislation-dependent resources for agency users (IRM and public information office personnel), Congressional users, and public policy researchers.

C3.6 Agency Program (mandatory if AIS)
Findings: The following tables present utilization of this element across the entire sample population because "AIS" as a record type could not be determined (see P4-Record Type). More than 70% of sampled records utilized this element. While time constraints precluded a systematic analysis of Agency Program values, the investigators noted only one record's reference to "OMB Circular A-130."

TOTAL SAMPLE

AGENCY PROGRAM (AIS)	N	%
YES	63	76%
NO	20	24%
TOTAL	83	100%

CORE SUBSET

AGENCY PROGRAM (AIS)	N	%
YES	33	79%
NO	9	21%
TOTAL	42	100%

Discussion: See P5-Information Object; given a liberal interpretation that AISs comprise "subject-matter databases," "agency homepages," "bibliographic databases," and "systems of systems," one may extrapolate only a predicted 50% utilization of this element. It is possible that "non-AIS resource" record creators are recognizing the instrinsic value of this element in educating public users about agency functions. NARA *Guidelines* state that "This element identifies the major agency program or mission supported by the system and should include a citation for any specific legislative authorities associated with this information resource…In general terms, it explains why the information resource was created in the first place. The rationale for a specific design is found in the PURPOSE element."

Recommendations: The rationale for isolating mandatory use of this element to describe AISs as well as differentiation of expected values between Purpose and Agency Program should be clarified in the NARA *Guidelines*. See also C3.5-Purpose recommendation concerning legislative citation.

C3.7 Availability-Distributor

Findings: 90% of records examined utilized at least one subordinate field of the Availability- Distributor subelement. Incidence of use in the Core subset was identical to that of the total sample.

TOTAL SAMPLE

AVAILABILITY-DISTRIBUTOR	N	%
YES	77	93%
NO	6	7%
TOTAL	83	100%

CORE SUBSET

AVAILABILITY-DISTRIBUTOR	N	%
YES	39	93%
NO	3	7%
TOTAL	42	100%

Discussion: Per the NARA *Guidelines*, the mandatory Availability element "is a grouping of subelements that together describe how the information resource is made available." Instructions for the mandatory Distributor subelement state: "Complete as many of the subordinate fields as necessary to identify the party from whom the information resource is available." Subordinate fields are Distributor's Name, Organization, Street Address, City, State, Zip Code, Country, Network Address, Hours of Service, Telephone, and Fax. The word choice "necessary to identify to the party" may be too vague to promote standard usage.

Recommendations: The *Guidelines* wording "Complete as many of the subordinate fields as necessary to identify the party" could be improved to "Complete all subordinate fields as available to assist the user in communicating with the Distributor."

C3.8 Availability-Order Process

Findings: Utilization (86%) of at least one subordinate field of the Availability- Distributor subelement was identical for the total sample and Core subset. One agency's sampled records contained a request for users to specify that the order is "in response to information in a…GILS record."

TOTAL SAMPLE

AVAILABILITY-ORDER PROCESS	N	%
YES	71	86%
NO	12	14%
TOTAL	83	100%

CORE SUBSET

AVAILABILITY-ORDER PROCESS	N	%
YES	36	86%
NO	6	14%
TOTAL	42	100%

Discussion: Per the NARA *Guidelines*, the mandatory Availability element "is a grouping of subelements that together describe how the information resource is made available." Instructions for the mandatory Order Process subelement state:

> Provide information on the common ways in which copies of the information resource or data from the resource may be obtained from this distributor. Alternatively, provide information on how the resource may be accessed. Note here if there is an electronic linkage for ordering, and also complete the LINKAGE subelement. Each agency may wish to establish standard ordering instructions for inclusion here.

The 86% utilization rate, compared with those of other mandatory elements and subelements, may be relatively low as a result of record creators entering "ordering" information in other elements [e.g., Access Constraints (see C3.10) or Linkage as noted above] or omission of this element altogether for nonaccessible resources (e.g., proprietary databases) or those that can not be "ordered" per se, such as an agency library.

Recommendations: Content analysis of the *Guidelines* by a third-party (i.e., not someone who was involved with the writing) is recommended to reveal areas of redundancy and ambiguity. Another approach would involve isolation of a random sample of selected data values (e.g., a distributor's URL) and testing whether GILS-cognizant vs. noncognizant users can place them in the "correct" (per published standards) elements.

C3.9 Sources of Data (mandatory if AIS)

Findings: The following tables present utilization of this element across the entire sample population because "AIS" as a record type could not be determined (see P4-Record Type). More than 50% of the records in the total sample utilized this element; utilization was lower (43%) in the Core subset.

TOTAL SAMPLE

SOURCE OF DATA (AIS)	N	%
YES	42	51%
NO	41	49%
TOTAL	83	100%

CORE SUBSET

SOURCE OF DATA (AIS)	N	%
YES	18	43%
NO	24	57%
TOTAL	42	100%

Discussion: See P5-Information Object; given a liberal interpretation that AISs comprise "subject-matter databases," "agency homepages," "bibliographic databases," and "systems of systems," one may extrapolate the found 50% utilization of this element. However, although values of this element were not analyzed systematically, the investigators noted several instances of non-AIS usage (e.g., the record "Detailed Briefing Materials," which describes a print supplement to the agency's Annual Budget press release, cites the office responsible). NARA *Guidelines* instruct: "Give information about the primary sources or providers of data to the system. State if the information in the system is generated by the agency, or if it is received by the system from outside the agency. Examples of the source of information for a system from outside the agency include corporations doing business in the U.S., broadcast license holders, or another Federal agency. This practice possibly indicates that record creators recognize the intrinsic value of this element in educating public users about agency methods and procedures for collecting information.

Recommendations: The rationale for isolating mandatory use of this element to describe AISs should be clarified in the NARA *Guidelines*.

C3.10   Access Constraints

Finding: This element was utilized in 99% of the total sample's records and in 100% of Core subset's records. The tables below reveal a slight difference between samples' incidence of "none" (the NARA-recommended null value expression) vs. substantive values. More than 25% of the sampled records describe a resource with at least one access constraint.

TOTAL SAMPLE

ACCESS CONSTRAINT	N	%
"NONE"	58	70%
SUBSTANTIVE	24	29%
NOT USED	1	1%
TOTAL	83	100%

CORE SUBSET

ACCESS CONSTRAINT	N	%
"NONE"	33	79%
SUBSTANTIVE	9	21%
NOT USED	0	0%
TOTAL	42	100%

Discussion: See C3.8-Order Process concerning possible redundancy of NARA-recommended Access Constraint values; *Guidelines* instructions for Access Constraint state to "[include] other special restrictions or limitations on obtaining the information resource. Guidance on obtaining any users' manuals or other aids needed for the public to reasonably access the information resource must also be included here [in the Documentation subelement]." (The latter direction raises an additional issue of whether recommended "users' manuals or other aids" are to be described by separate but complementary GILS record(s) and, if so, whether the at-hand GILS record's Cross-Reference value or Documentation value should link to the actual manual or to the GILS record describing it. GPO's GILS records entitled "GPO Access" (Control Identifier: "GPO/SOD/OEIDS00027" and "GPO Access User Guide Online via GPO Access" (Control Identifier: "GPO/SOD/OEIDS-00026")) provide an example of this scenario. In addition, see discussion at C3.11-Use Constraints for possible confusion over "access" vs. "use" constraint elements' values.

Recommendations: Investigators recommend that the NARA *Guidelines* differentiate placement of certain constraints (e.g., "requires forms-capable browser"—in Access Constraint vs. Order Process) and clarify the relationship between or the presence/absence of a user's manual or aid and "access constraints."

C3.11　　Use Constraints

Findings: This element was utilized in 100% of both the total sample's and Core subset's records. The tables below reveal a slight difference between samples' incidence of "none" (the NARA-recommended null value expression) vs. substantive values.

TOTAL SAMPLE

USE CONSTRAINT	N	%
"NONE"	69	83%
NAMED	14	17%
TOTAL	83	100%

CORE SUBSET

USE CONSTRAINT	N	%
"NONE"	34	81%
NAMED	8	19%
TOTAL	42	100%

Discussion: While analysis of the substantive values in this element was not within the scope of the current study, the investigators noted several instances of Use Constraint values referring to Access Constraints (e.g., "[agency service organization] does not lend reference materials…although the public is welcome to use all of these in-house"). One record populated both the access and use constraint elements with "This system is an internal information and processing system and is not generally available for review outside the agency." Use Constraints is defined by the NARA *Guidelines* as: "describ[ing] any constraints or legal prerequisites for using the information resource or its component products or services." The definition of Access Constraints (see C3.10 above) carries precisely the same wording save for the substitution of the word "accessing" for "using." However, NARA's usage guidelines, which contain differentiating caveats (e.g., "Restrictions on what may be done with the information once it has been accessed are found in the USE CONSTRAINTS data element"), for both these elements appear to have been missed by some record creators.

Recommendation: Upon confirmation that confusion among Access vs. Use Constraints is significant, it is recommended that Use Constraints be renamed "Restrictions on Use of Information" or some other more straightforward phrase.

C3.12　　Point of Contact

Findings: Assessment of utilization for this element considered a substantive value in any subelement. 99% of the total sample and 98% of the Core subset records utilized this element. "Point of Contact Type" below reflects the initial subelement value in this element. 1 in 4 records sampled featured a personal name; roughly half featured the name of an office. 20% of the values were a job title.

TOTAL SAMPLE

POINT OF CONTACT TYPE	N	%
NAME	19	23%
OFFICE	45	54%
JOB TITLE	16	19%
OTHER	2	2%
NOT USED	1	1%
TOTAL	83	100%

CORE SUBSET

POINT OF CONTACT TYPE	N	%
NAME	7	17%
OFFICE	24	57%
JOB TITLE	8	19%
OTHER	2	5%
NOT USED	1	2%
TOTAL	42	100%

Discussion: Per the NARA *Guidelines*, this element "identifies an organization, and a person where appropriate, serving as the point of contact plus methods that may be used to make contact. This element consists of the following subelements: [Name; Organization; Street Address; City; State; Zip Code; Country; Network Address; Hours of Service; Telephone; Fax]…Complete as many of the subordinate fields as are necessary to identify the organization and individual *responsible for the content of the information dissemination product or automated information system* [investigators' emphasis]. While confirming the accuracy of Points of Contact was beyond the scope of the current analysis, it is doubtful that "authors" are being listed in this element in cases where discrete information products (see P5-Objects Represented) are being described.

Recommendation: It is suggested that research be conducted to assess users' expectations concerning presentation of true "authorship" data in the bibliographic sense as opposed to the entity responsible for compilation, administrative maintenance, or dissemination of the resource—a value they may expect feasibly to find in Originator (whose definition includes "organizational unit that *created and maintains* [investigators' emphasis] the information dissemination product or information system" and Record Source elements as well. In addition, continuous and unscheduled audit of the accuracy of Points of Contact, as well as the quality of Contacts' responses to GILS record-related questions, is essential. Users may not only abandon GILS as a result of a discrepancy in this element but broadcast the failure among communities of interest.

C3.13   Schedule Number (mandatory if intended to meet the obligation…to inventory automated information systems or other records series for records management purposes)

Findings: The following tables present incidence of populated Schedule Number elements and their respective values. For the total sample, 55% of records utilized this element. 14% of all records sampled contained a record schedule number; 41% contained values indicating that scheduling was not required, is pending, etc. For the Core subset, an identical 55% element utilization rate was found. 5% of Core records contained a schedule number and 50% contained values indicating that scheduling was not required, is pending, etc.

TOTAL SAMPLE

SCHEDULE NUMBER	N	%
NOT USED (TOTAL)	37	45%
"NOT SCHEDULED"	16	
"N/A"	4	
"PENDING"	3	
"NONE"	3	
"NOT APPLICABLE"	2	
"UNSCHEDULED"	2	
"SCHEDULE IN PROGRESS"	2	
THIS IS A NONRECORD INFORMATION DISSEMINATION PRODUCT	1	
SCHEDULE TO BE SUBMITTED PENDING NARA INFORMAL REVIEW OF . . . PUBLICATIONS	1	
TOTAL	34	41%
ATF RCS 201, ITEM 140	1	
GENERAL RECORDS SCHEDULE NUMBER #20.9	1	
GRS 14 Sec. 6 and GRS 20 Sec. 9	1	
N1-138-88-2	1	
N1-266-77-2-92	1	
N1-309-87-002	1	
N1-522-95-1	1	
N1-95-88-2/62-9.11	1	
NC1-122-79-1, ITEM 11	1	
PBGC ITEM #67	1	
SCHEDULED- N1420-93-1, #26	1	
SCHEDULED-N1-420-93-1, #12	1	
TOTAL	12	14%
GRAND TOTAL	83	100%

CORE SUBSET

SCHEDULE NUMBER	N	%
NOT USED (TOTAL)	19	45%
"NOT SCHEDULED"	11	
"N/A"	0	
"PENDING"	3	
"NONE"	3	
"NOT APPLICABLE"	2	
"UNSCHEDULED"	0	
"SCHEDULE IN PROGRESS"	2	
THIS IS A NONRECORD INFORMATION DISSEMINATION PRODUCT	0	
SCHEDULE TO BE SUBMITTED PENDING NARA INFORMAL REVIEW OF . . . PUBLICATIONS	0	
TOTAL	21	50%
N1-266-77-2-92	1	
N1-95-88-2/62-9.11	1	
TOTAL	2	5%
GRAND TOTAL	42	100%

Discussion and Recommendations: See Chapter 4 for discussion of GILS and records management.

C3.14   Control Identifier

Findings: This element was utilized in the total sample and Core subsets at 91% and 88%, respectively.

TOTAL SAMPLE

CONTROL ID	N	%
YES	76	91%
NO	7	9%
TOTAL	83	100%

CORE SUBSET

CONTROL ID	N	%
YES	37	88%
NO	5	2%
TOTAL	42	100%

Discussion: NARA *Guidelines* state that "this element is defined by the *information provider* [investigators' emphasis] and is used to distinguish this *locator record* [investigators' emphasis] from all other GILS Core entries. The control identifier should be distinguished with the record source agency acronym as provided in the *U.S. Government Manual*….Create a unique identifying number for each GILS Core entry. The control identifier will consist of two parts: an identifying acronym followed by a control number." While no explanation of the lower utilization rate for the Core subset can be offered by the investigators, evidence was found of two possibly related problems.

First, college-educated participants in the online user assessment (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations) did not understand the definition or use of this element. Upon presentation of the term only, prior to any searching, 8 of 10 stated they were "uncomfortable with, or unsure of using" the element; upon presentation of the NARA definition, 4 of 8 respondents stated that "the definition, and how it fits into GILS" was still "unclear." It may not be unreasonable to assume that the users' experience is transferable to a record-creator.

Second, "information provider" is an undefined but crucial term within the element's definition and could cause confusion when opposed to the construct of "author" or distributor (see further discussion of this issue at C3.12-Point of Contact). In addition, the definition assumes an understanding of the concept of "core" and "locator" records that may be unwarranted even among those in the GILS community—given the overall utilization rate of mandatory elements by the Core subset compared with the total sample. Finally, the *Guidelines* instructions concerning the values for this element *vis-a-vis* the many if/then scenarios are relatively elaborate and may require a level of inter-agency cooperation not in evidence during this study.

An assessment of use of agency acronyms as set forth in the *U.S. Government Manual* was not performed due to time constraints, however it is noted that this requirement will not serve the stated purpose of supporting NIDR unless the user has ready access to the *Government Manual*.

Recommendations: Further assessments of utilization rates, including an accounting for the quality of values as well as the rationale used in their construction/assignment appears warranted. In addition, on GPO Access GILS, it may be prudent to provide a hypertext link from "*US Government Manual*" in the field definition files to a recast version of *Government Manual* Appendix A: Commonly Used Abbreviations and Acronyms.

C3.15   Record Source

Findings: 90% of the total sample's records utilized this element; 88% of the Core subset records did. 63% of the records sampled utilized at least one Record Source subelement but only 46% presented its label.

TOTAL SAMPLE

RECORD SOURCE	N	%
YES	74	90%
NO	9	10%
TOTAL	83	100%

CORE SUBSET

RECORD SOURCE	N	%
YES	37	88%
NO	5	2%
TOTAL	42	100%

TOTAL SAMPLE

RECORD SOURCE SUBELEMENTS	N	%
YES	52	63%
NO	31	37%
TOTAL	83	100%

CORE SUBSET

RECORD SOURCE SUBELEMENTS	N	%
YES	27	64%
NO	15	36%
TOTAL	42	100%

TOTAL SAMPLE

RECORD SOURCE SUBELEMENTS LABELED	N	%
YES	38	46%
NO	45	54%
TOTAL	83	100%

CORE SUBSET

RECORD SOURCE SUBELEMENTS LABELED	N	%
YES	16	38%
NO	26	62%
TOTAL	42	100%

Discussion: "This element identifies the organization, as named in the *U.S. Government Manual*, that created or last modified this locator record…Give the name of an organization, and normally the name of the unit, that has created this GILS Core entry" per the NARA *Guidelines*. Most study participants felt that "all GILS records should look alike" (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations), and it may be concluded that this preference refers to the presence and absence of display characteristics rather than content. Investigators believe that standardization of element label display will contribute to users' "footing" within GILS vs. other information space—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. An assessment of use of agency names as set forth in the *U.S. Government Manual* was not performed due to time constraints, however it is noted that this requirement will not serve the stated purpose of supporting NIDR unless the user has ready access to the *Government Manual* Table of Contents or Appendix A: Commonly Used Abbreviations and Acronyms.

Recommendation: Further assessment of user preferences for display of organizational subelement labels and values, as well as their aggregation levels and resultant maintenance burden, is recommended. In addition, on GPO Access GILS, it may be prudent to provide a hypertext link from "*US Government Manual*" in the field definition files to a recast version of *Government Manual* Appendix A: Commonly Used Abbreviations and Acronyms.

C3.16 Date of Last Modification

Findings: Utilization of this element was identical in the total and Core subset samples: 98%. The tables below also reveal a high +33% incidence of nonstandard date formats in sampled GILS records (where *N*=1 instance of nonstandard date, in *any* element). In addition, the data reveal that at least four sampled records contained values indicating a misperception that the value of the element refers to the information resource being described rather than the record itself: "FY 1994," "1988," "1989," and "continuously updated."

TOTAL SAMPLE

DATE OF LAST MOD	N	%
YES	81	98%
NO	2	2%
TOTAL	83	100%

CORE SUBSET

DATE OF LAST MOD	N	%
YES	41	98%
NO	1	2%
TOTAL	42	100%

TOTAL SAMPLE

LAST MOD DATE	
NOT USED (N)	2
MOST RECENT	THURSDAY, DECEMBER 12, 1996
EARLIEST	SATURDAY, DECEMBER 31, 1988
MODE	TUESDAY, DECEMBER 12, 1995
AVERAGE	SUNDAY, DECEMBER 31, 1995

CORE SUBSET

LAST MOD DATE	
NOT USED (N)	2
MOST RECENT	THURSDAY, DECEMBER 12, 1996
EARLIEST	MONDAY, DECEMBER 12, 1994
MODE	TUESDAY, DECEMBER 12, 1995
AVERAGE	SATURDAY, MARCH 16, 1996

TOTAL SAMPLE

ALL DATES IN YYYYMMDD?	N	%
YES	51	61%
NO	32	39%
TOTAL	83	100%

CORE SUBSET

ALL DATES IN YYYYMMDD?	N	%
YES	25	60%
NO	17	40%
TOTAL	42	100%

Discussion: The NARA *Guidelines* definition of this element is: "This element identifies the latest date on which this locator record was created or modified"—a relatively straightforward statement that is unfortunately obfuscated by the following [investigators' emphasis]:

> Usage Guidelines: This element is used initially to record the date of the creation of the GILS Core *entry*. The value of the element changes with any subsequent modifications [*to what?*]. Only the date of the last modification of the *entry* needs be included in the GILS Core *entry*, but agencies may wish to track the date of the initial creation of the GILS *record* as well in a local [sic] defined element.

An interview with GPO Access GILS support staff corroborated record-creators' confusion concerning whether the value of this element describes the information resource or the GILS record itself. In addition, more than half of participants in the scripted online user assessment chose the value appearing in this element over others when asked to supply "how up-to-date are the described materials [in the record]?"

Recommendations: It is highly recommended that the Date of Last Modification element be renamed "Record Revision Date" and the confusing "entry" nomenclature in the usage guidelines be eliminated. In addition, where software resources permit, "auto-correct/auto-format" macros should be utilized to standardize dates. Use of computer-generated time/stamp dates, perhaps as a final "record checked and approved for release" procedure, may promote responsibility for record content quality among creators, facilitate audit sampling, and ease record maintenance as well. A possible future research project could assess government-wide frequency of change and clustering of Dates of Last Modification to reveal patterns in record maintenance.

C4   Utilization and Characteristics of Selected "Optional" Elements

The following list summarizes findings from this portion of the analysis.

- Utilization of Controlled Vocabulary was assessed at 10%, with *Library of Congress Subject Headings* being the most often cited Thesaurus
- About half the records sampled featured at least one Local Subject Index term—nearly all being a variant of "US Federal GILS"
- Resource Description was provided in only 12% of the total sample but 24% of the Core subset
- Utilization of Methodology was negligible.

The relatively low incidence of nonmandatory elements in the sampled records is not seen as indicative of their potential application—i.e., nothing about the nature of the information resources described was revealed as precluding or discouraging their use. Rather it may reflect a minimum-compliance or expediency-based approach to record creation, the degree of comprehensibility of or intellectual accessibility to the NARA *Guidelines*, and/or insufficient training as to the value of elements in networked information discovery and retrieval (NIDR).

The following results and discussion are presented per element in "preferred display" order for a GILS record (*FIPS Pub. 192* and NARA *Guidelines*). Note that the term utilization here means that the element was present, presented, and populated (even in cases where "none", "n/a", or similar acknowledgment of an absence of substantive content was present).

C4.1 Controlled Vocabulary-Index Terms-Controlled
Findings: 10% percent of sampled records utilized controlled index terms; for the Core subset, utilization was slightly higher.

TOTAL SAMPLE

INDEX TERMS-CONTROLLED	N	%
YES	8	10%
NO	75	90%
TOTAL	83	100%

CORE SUBSET

INDEX TERMS-CONTROLLED	N	%
YES	5	12%
NO	37	88%
TOTAL	42	100%

Discussion: The NARA *Guidelines* state the benefits of this element very well: "One method of identifying possible GILS entries of interest will be provided through the ability in GILS to search the entire text of the entry, including the narrative description in the ABSTRACT element. More precise search results can be achieved through the use of the CONTROLLED VOCABULARY element." There is no doubt that inclusion of controlled vocabulary terms greatly enhances NIDR. In addition, use of registered thesauri (see C4.2-Controlled Vocabulary-Thesaurus) may orient users in the resource subject domain; the majority of participants in the online assessment of GILS believed "It would be easier to search GILS records if they were grouped hierarchically by subject." (See Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations).
Recommendations: It is recommended that agencies pursue research into the effects of "familiar" and specialized Controlled Vocabularies on NIDR, in concert with the Library of Congress and GPO's Cataloging Branch, to lessen users' dependency on knowledge of agency mission and to increase precision of information retrieval. In addition, the value of providing thesaurus hyperlinkages warrants systematic study as does user preferences for "catalog-based" NIDR (e.g., the approach of Yahoo! and the Argus Clearinghouse). Users who find relevant records easily will probably focus less on GILS "user-ugliness" and serve as the most effective mode of promotion.

C4.2 Controlled Vocabulary-Thesaurus

Findings: Utilization of this element was higher for the Core subset (17%). The total sample named a thesaurus 8% of the time. *Library of Congress Subject Headings* was the most frequent value.

TOTAL SAMPLE

THESAURI	N	%
HAZARDOUS WASTER SUPERFUND DATABASE	2	2%
LCSH	4	5%
DTIC	1	1%
TOTAL	7	8%

CORE SUBSET

THESAURI	N	%
HAZARDOUS WASTER SUPERFUND DATABASE	2	5%
LCSH	4	10%
DTIC	1	2%
TOTAL	7	17%

Discussion and Recommendations: See C4.1-Controlled Vocabulary-Index Terms-Controlled.

C4.3 Local Subject Index

Findings: This element was used in 98% of the core subset (see note at data), and in more than half of the total sample.

TOTAL SAMPLE

LOCAL SUBJECT TERMS	N	%
YES	45	54%
NO	38	46%
TOTAL	83	100%

CORE SUBSET

LOCAL SUBJECT TERMS	N	%
YES	41	98%
NO*	1	2%
TOTAL	42	100%

*This results from the appearance of "US Federal GILS" in an element other than Local Subject Index.

Discussion: The utilization rate of this element within the Core subset is largely attributable to the minimal inclusion of the required "US Federal GILS" or variant. The NARA *Guidelines* define usage of the element for: "supplement[ing] an existing thesauri or in the absence of an acceptable listing…Identify significant subjects that apply to the information resource including broad concepts and unusual aspects of the system or product. Use topical subject headings consisting of general subject terms or names of events or objects."
Recommendation: An assessment of the frequency of use of Local Subject Terms other than to denominate a "core" record is highly recommended as a first step for clarifying the usefulness of this element in public-access NIDR. It is possible that this element will evolve to describe "aspects of the system or product" and/or "names of events or objects" relevant to Internet information space navigation, for example: "gopher archive, listserv, SIG" or "census, PDF, Web download"; see also C4.4-Availability-Resource Description below.

C4.4 Availability-Resource Description

Findings: Only 12% the total sample's records utilized this element; utilization was double in the Core subset.

TOTAL SAMPLE

AVAILABILITY-RESOURCE DESCRIPTION	N	%
NO	73	88%
YES	10	12%
TOTAL	83	100%

CORE SUBSET

AVAILABILITY-RESOURCE DESCRIPTION	N	%
NO	32	76%
YES	10	24%
TOTAL	42	100%

Discussion: The NARA *Guidelines* read in part: "This subelement identifies the resource as it is known to the distributor." Difficulties encountered while characterizing GILS information resources (see P4 Record Types and P5-Objects Represented), aggregation levels (see P6-Aggregation), and dissemination media (see P7-Containers)

indicate that this element (or an additional element with this name) might better serve the objective of public-access NIDR if it were redefined to comprise the object/aggregation/container concept.

Recommendations: A qualitative analysis of element values for a large sample of GILS records is highly recommended. (See P4-Record Types, P5-Objects Represented, P6-Aggregation, and P7-Containers.)

C4.5 Methodology

Findings: Utilization of this element was negligible (2%) in the total sample; the Core subset reflected 5% utilization. One record's value for this element contained information about the manufacturing process for a CD-ROM (as opposed to the definitional content that would have described how the data on it were collected/compiled).

TOTAL SAMPLE

METHODOLOGY	N	%
NO	81	98%
YES	2	2%
TOTAL	83	100%

CORE SUBSET

METHODOLOGY	N	%
NO	40	95%
YES	2	5%
TOTAL	42	100%

Discussion: These findings may indicate another instance confusing application instructions. The NARA *Guidelines* state that the Methodology element "identifies any specialized tools, techniques, or methodology used to *produce* [investigators' emphasis] this information resource...Provide here information concerning significant methodological characteristics of the information resource. Examples of items that might be discussed include the algorithm, universe description, sampling procedures, classification, or validation characteristics." An incomplete reading of instructions might lead to the CD-ROM example cited above, and a record-creator's lack of knowledge of research terminology or access to relevant information might prohibit inclusion of a methodology description altogether.

Recommendations: This element, like Sources of Data (see C3.9) for AISs in particular, demands the contribution of resource creators or collaborators. Because its absence may subject the information resource to misuse or abandonment, it is recommended that agencies develop training and complementary procedures for record creators to recognize the need for and to obtain the information for populating this element.

4.3.   Resource Profile

Appendix C-4 Record Content Analysis Methodology enumerates the 42 agencies whose records were included in the record content analysis. The following paragraphs provide chief defining characteristics of the sample.

Subject-matter databases (e.g., that for red cockaded woodpecker or accident-investigation information) comprised more that 20% of the resources described, followed by discrete publications (19%) and miscellaneous documents in an ad hoc collection (17%) (see P5-Objects Represented and Appendix C-4 Record Content Analysis Methodology for semantics). Agency homepages comprised 10% of the sampled records, a positive indicator of NIDR synergy among Federal information spaces. An analysis of aggregation types (see P6-Record Aggregation and Appendix C-4 Record Content Analysis Methodology for semantics) revealed that more than one-third of records described "new collections"—i.e., resources aggregated by virtue of the record itself (although the serviceability of this phenomenon in NIDR requires additional study). Related to description of resource types and aggregation, dissemination media or "containers" of information were found to be largely unidentifiable by the record's content; 22% of resources were described as being packaged/disseminated by multiple modes, and the highest identifiable single mode was print (23%) followed by Web (8%).

The most important finding of this "profile" analysis is that GILS covers a potentially unfathomable scope of information, and that the implications of this for record creation and NIDR demand consideration and commentary by policymakers. The resolution of issues such as:

- Optimal aggregation (i.e., unit of resource selected for representation)
- Enumeration and description of an increasing multiplicity of containers (and the potential for resultant resource derivation, amendment, or abridgment and ensuing authority, integrity, and maintenance concerns)

- Determination of how GILS should be presented to maximize users orientation in Federal information space *vis-a-vis* cross-agency search capabilities and synergy with agencies other metadata and full-text resources)

are perceived as fundamental to GILS ultimate success.

P1   Record Types
Findings: The investigators was unable to code sampled records according to record types of Automated Information System (AIS), Locator, and Privacy Act systems as outlined in the NARA *Guidelines*.
Discussion: (See also P5-Objects Represented.) The NARA *Guidelines* delineates three types of information resources: locators, automated information systems (AISs), and Privacy Act systems of records. The following text is provided to highlight the burden of denominating record types.

The *Guidelines* define "locators" by means of example:

…locators (not to be confused with the GILS itself) catalog or describe information products (such as books, CD-ROMs, publications, studies, reports, and patents, regardless of medium). For example, there could be a GILS Core entry describing the Library of Congress Information System (LOCIS). LOCIS consists of catalog entries for publications, and hence, serves as a locator to these publications

and "AISs" by means of reference to OMB Circular No. A-130:

An automated information system is a discrete set of information resources organized using information technology as defined in OMB Circular No. A-130

which, in turn, provides no definition of an AIS *per se* but defines "information resource" by way of inclusion:

The term "information resource" includes both government information and information technology

where:

The term "Information technology" means the hardware and software operated by a Federal agency…to accomplish a Federal function

as well as by functionality:

[AISs] may be used for the collection, processing, maintenance *or* [investigators' emphasis] dissemination of information, including Federal records…

A "Privacy Act system of records," the third information type delineated in the NARA *Guidelines*, is therein defined by way of reference to U.S.C. 552a, which specifies:

the term 'system of records' means a group of any records under the control of any agency from which information is retrieved by…[an] identifying particular assigned to the individual.

This brief definitional synopsis is presented to stress the difficulty, and perhaps futility, of classifying GILS records by the information types set forth in the NARA *Guidelines*. The definitions are not mutually exclusive (i.e., many "locators" and "Privacy Act systems" are also "AISs") and the investigators felt that intercoder reliability was sufficiently poor to abandon the task. However, it is noted that five records sampled specified "Privacy Act system."

The researchers found that a code of "object represented," developed during the current record content analysis, proved both more straightforward and productive in characterizing GILS records' content. Results of this approach are discussed in the following section.

P2   Objects Represented

Findings: This analysis aimed at capturing what types of information products and resources GILS records describe in accordance with the categories shown in Table E2-2.

Table E2-2
Aggregation Semantics

Code	Operational Definition	Examples
Record Aggregates Objects	GILS record, by virtue of its creation, collects discrete information resources that record content indicates would not have otherwise been collected or aggregated. Assigned in the absence of clues within the record that the represented objects were heretofore packaged *as this collection* to optimize information discovery and retrieval.	- Privacy Act Systems compilation - files - press releases - forms
Aggregated Object Represented	GILS record represents an *a priori* or purposeful collection of information resources—e.g., woodpecker database or agency website. GILS record represents an object that collects, or comprises, two or more discrete information objects, and that represents a collection of standalone information files or products packaged together on the basis of a common theme or subject for functional convenience.	- CD-ROM of regulations - System that compiles Privacy Act records - job line of open requisitions
Discrete Object Represented	GILS record describes a standalone document-level entity that does not meet the criteria for "object aggregates metadata" below.	- annual report - videotape
Object Aggregates Metadata	GILS record describes a pre-existing metadata collection, or "locator," as an information resource.	- directory - catalog - index - log

As shown in the following data summaries, "subject matter databases" (i.e., a database of endangered species) were often described, comprising one-quarter of the records sampled. The Core subset sample reflected a higher value for "publication"—a category including discrete information objects available as self-contained entities such as books and individual technical reports, etc.

TOTAL SAMPLE

OBJECT REPRESENTED	N	%
SUBJECT MATTER DATABASE	18	22%
PUBLICATION	16	19%
MISCELLANEOUS DOCUMENTS IN *AD HOC* COLLECTION	14	17%
AGENCY HOMEPAGE	8	10%
ORGANIZATION	6	7%
FORM	4	5%
ADMINISTRATIVE CATALOG	3	4%
BIBLIOGRAPHIC DATABASE	3	4%
PUBLICATIONS CATALOG	4	5%
SYSTEM OF SYSTEMS	3	4%
PROGRAM	2	2%
JOB LINE	1	1%
UNKNOWN	1	1%
TOTAL	83	100%

CORE SUBSET

OBJECT REPRESENTED	N	%
PUBLICATION	11	26%
SUBJECT MATTER DATABASE	8	19%
MISCELLANEOUS DOCUMENTS IN *AD HOC* COLLECTION	7	17%
AGENCY HOMEPAGE	4	10%
ORGANIZATION	2	5%
FORM	2	5%
PUBLICATIONS CATALOG	2	5%
SYSTEM OF SYSTEMS	2	5%
ADMINISTRATIVE CATALOG	1	2%
BIBLIOGRAPHIC DATABASE	1	2%
PROGRAM	1	2%
JOB LINE	1	2%
UNKNOWN	0	0%
TOTAL	42	100%

Discussion and Recommendations: The findings above show that nearly 20% of records sampled described a discrete publication as opposed to a "locator" (e.g., catalog) resource. This perhaps indicates an incorrect assumption by policymakers that all of an agency's "information dissemination products" (including any book, paper, map, machine-readable material, audiovisual production, or other documentary material, regardless of physical form or characteristic, disseminated by an agency to the public") were described by extant locators. In fact, "locator-function" objects (extrapolated from Administrative Catalog, Bibliographic Database, and Publications Catalog) represented only a little more than 10% of objects described in the sample. See C3.1-Title, C3.4-Abstract, C3.6-Agency Program, C4.4-Availability-Resource Description, and P6-Aggregation for implications and recommendations associated with identifying "objects" within GILS records. In summary, it is highly recommended that users be provided an additional GILS element of "object represented" in order to evaluate aggregation, or "informational distance" from satisfaction of their requirement, and to increase retrieval precision (e.g., by eliminating object type "job line" from a search on unemployment statistics).

P3   Record Aggregation

Findings: The analysis attempted to characterize each sampled record's "aggregation" or "granularity" relative to others in the sample. Appendix C-4 Record Content Analysis Methodology provides a complete discussion of semantics; however, a summary of definitions is provided below.

- *Record aggregates object*. The GILS record, by virtue of its creation, collects discrete information resources that the record content indicates would not have otherwise been collected or aggregated—e.g., "Privacy Act Systems," "General Files", "Press Releases", or "Forms". This code was assigned in the absence of clues within the record that the represented objects were heretofore packaged *as this collection* to optimize information discovery and retrieval.
- *Aggregated object represented*. The GILS record represents an *a priori* or purposeful collection of information resources—e.g., "Woodpecker Database" or an agency Web site. In other words, the GILS record represents an object that collects, or comprises, two or more discrete information objects, and that object represents a collection of standalone information files or products packaged together on the basis of a common theme or subject for functional convenience—e.g., a CD-ROM of regulations, a discrete system of Privacy Act records, and a voice recording of employment opportunities.
- *Discrete object represented*. The GILS record describes a standalone document-level entity that does not meet the criteria for "object aggregates metadata" below—e.g., an Annual Report or videotape.
- *Object aggregates metadata*. The GILS record describes a pre-existing metadata collection, or "locator"—e.g., directory, catalog, index, or log—as an information resource.

The analysis revealed a high number of records, more than one-third of the sample, that appeared to aggregate records, document, files, and other objects. Attachment 2a, a record titled "Briefing Materials, Public Comments, Other Related Official Files" exemplifies this phenomenon.

Records describing a discrete information resource that comprises two or more sub-resources (aggregated object represented) constituted 25% and 31% of the total and Core subset sample, respectively. Attachment 2b, a record titled "Worldwide Real Property Inventory System," describes a system that aggregates data from other systems.

Approximately one-fourth of the records sampled described a discrete, standalone information object such as a book, video, or technical report (see Attachment 2c "Investment Fund Brochure"). Metadata collections were described about 1 in 10 records; Attachment 2d, a record titled "Office of the General Counsel Library Catalog," serves as an example of items in this category. Five records whose "objects represented" (see P5) were not information resources (e.g., "organizations," "[functional] programs," and "facilities") could not be coded as to aggregation.

TOTAL SAMPLE

AGGREGATION	N	%
RECORD AGGREGATES OBJECTS	30	36%
AGGREGATED OBJECT REPRESENTED	21	25%
DISCRETE OBJECT REPRESENTED	17	20%
OBJECT AGGREGATES METADATA	10	12%
UNKNOWN	5	6%
TOTAL	83	100%

CORE SUBSET

AGGREGATION	N	%
RECORD AGGREGATES OBJECTS	16	38%
AGGREGATED OBJECT REPRESENTED	13	31%
DISCRETE OBJECT REPRESENTED	7	17%
OBJECT AGGREGATES METADATA	4	10%
UNKNOWN	2	5%
TOTAL	42	100%

Discussion and Recommendations: More than one-third of records described "new collections"—i.e., resources aggregated by virtue of the record itself, although the serviceability of this phenomenon in NIDR is questionable given record titles such as "Minutes" and "General Files." Counterbalancing this finding, however, is a roughly 20% incidence of item-level description when both "record aggregates metadata" (e.g., a catalog is the object) and "discrete object represented" (e.g., a brochure is the object). Investigators found the task of characterizing granulation extremely challenging and recognize the problems of nonexclusivity and intercoder reliability in the method employed during this analysis. However, a similar coding scheme, perhaps based on clearly defined (including by way of example) steps of distance from satisfaction of an information need and deliberately associated with object type, may be useful in facilitating NIDR. Section C4.4-Availability-Resource Description addresses the issue of granularity as well.

P4   Containers

Findings: "Containers" were defined as dissemination media. Nearly half of the total sample described information resources whose containers could not be discerned from record content; containers in the Core subset were more frequently mentioned (31%). More than 1 in 4 of both the total and subset records described resources available in multiple containers. "Print" resources comprised about 20% of both samples, followed by "Web" resources (about 10%).

TOTAL SAMPLE

CONTAINER	N	%
UNKNOWN	34	41%
MULTIPLE (OF ALL)	18	22%
PRINT	19	23%
WEB	7	8%
DIALUP	2	2%
CD-ROM	1	1%
VOICE	1	1%
VIDEO	1	1%
RADIO/TV BROADCAST	0	0%
MICROFORM	0	0%
LISTSERVE	0	0%
GOPHER	0	0%
FTP	0	0%
FAX	0	0%
EMAIL	0	0%
TOTAL	83	100%

CORE SUBSET

CONTAINER	N	%
UNKNOWN	13	31%
MULTIPLE (OF ALL)	12	29%
PRINT	10	24%
WEB	4	10%
DIALUP	1	2%
CD-ROM	1	2%
VOICE	1	2%
VIDEO	0	0%
RADIO/TV BROADCAST	0	0%
MICROFORM	0	0%
LISTSERVE	0	0%
GOPHER	0	0%
FTP	0	0%
FAX	0	0%
EMAIL	0	0%
TOTAL	42	100%

Discussion: User recognition of and ability to access/use "containers" may be a significant factor in relevance judgment (see discussion at C4.4-Resource Description). While time constraints precluded an assessment of hypertext incidence within Web containers, S2-Content Hypertext provides data concerning total incidence within the sample.

Recommendations: See C4.4-Resource Description.

4.4.   Serviceability

The "serviceability" data shown below are considered to represent record effectiveness in terms of the degree to which they enhance NIDR, convenience to the user, aesthetics, readability, and relevance judgment.

S1   File Formats
Findings: The HTML file format was available for nearly half of the total records sampled and 60% of records in the Core subset. Two records were available in SGML; none were available in PDF.

TOTAL SAMPLE

ASCII TEXT AVAIL?	N	%	HTML AVAIL?	N	%	PDF AVAIL?	N	%	SGML AVAIL?	N	%
YES	67	81%	YES	38	46%	YES	0	0%	YES	2	2%
NO	16	19%	NO	45	54%	NO	83	100%	NO	81	98%
TOTAL	83	100%	TOTAL	83	100%	TOTAL	83	100%	TOTAL	83	100%

CORE SUBSET

ASCII TEXT AVAIL?	N	%	HTML AVAIL?	N	%	PDF AVAIL?	N	%	SGML AVAIL?	N	%
YES	32	76%	YES	25	60%	YES	0	0%	YES	2	5%
NO	10	24%	NO	17	40%	NO	42	100%	NO	40	95%
TOTAL	42	100%	TOTAL	42	100%	TOTAL	42	100%	TOTAL	42	100%

Discussion: Each file format has advantages for record creators and users, however given a choice between HTML- and ASCII-formatted records, study participants expressed a clear preference for the former (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations). See also A1-Format Errors. Also, it is noted that the use of the word "TEXT" vs. "HTML" on results lists may lead users to believe that clicking on the former will provide the (full) text of a document.

Recommendations: As noted in other sections, investigators believe that standardization of record display will contribute to users' "footing" within GILS vs. other information—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. The ready availability of "HTML editor" programs, which convert a variety of file formats to HTML, should be exploited as resources permit.

S2   Content Hypertext
Findings: Roughly 1 in 4 of the records sampled featured at least one instance of hypertext somewhere within element values. The following tables show that Available Linkage and Distributor Network address were the most frequently hotlinked elements (N = number of records). The incidence of Cross-Reference element hypertext was negligible.

TOTAL SAMPLE

ELEMENTS WITH HOTLINKED VALUE	N	% OF TOTAL USE
AVAILABLE LINKAGE	13	52%
DISTRIBUTOR NETWORK ADDRESS	6	24%
DISTRIBUTOR URL (LOCALLY DEFINED)	2	8%
ABSTRACT	1	4%
CROSS-REFERENCE LINKAGE	1	4%
CROSS-REFERENCE TITLE	1	4%
ORDER PROCESS	1	4%
TOTAL	25	100%
NUMBER OF RECORDS USING LINKS	19	23%

CORE SUBSET

ELEMENTS WITH HOTLINKED VALUE	N	% OF TOTAL USE
AVAILABLE LINKAGE	13	52%
DISTRIBUTOR NETWORK ADDRESS	6	24%
DISTRIBUTOR URL (LOCALLY DEFINED)	2	8%
ABSTRACT	1	4%
CROSS-REFERENCE LINKAGE	1	4%
CROSS-REFERENCE TITLE	1	4%
ORDER PROCESS	1	4%
TOTAL	25	100%
NUMBER OF RECORDS USING LINKS	12	29%

Discussion: While the maintenance burden of hypertext is recognized, users' expectation for it will continue to accelerate for the foreseeable future.

Recommendations: A further analysis of hypertext incidence and placement is warranted to capitalize on user expectations. In the Scripted Online User, for example, users were asked "What do you think would happen if you were to click on this record's hypertext title? ["U.S. International Trade Commission. Library Services"], and presented a multiple-choice list of:

    a.   I would jump to the ITC website
    b.   I would connect to ITC's online library catalog
    c.   A list of ITC library staff contacts would appear
    d.   I would link to a fuller/longer version of this record
    e.   I would be given a list of library services such as interlibrary loan, photocopying, and research assistance
    f.   Other (please specify)

It is recommended that system designers actively participate in and contribute to PURL and similar technology development efforts to ensure satisfaction of GILS unique requirements.

S3   Capitalization

Findings: 86% of records sampled used sentence-case capitalization; 10% used all capitals for element labels.

TOTAL SAMPLE

CAP STYLE	N	%
SENTENCE CASE	71	86%
ELEMENTS ONLY CAP	8	10%
MIXED	4	5%
OTHER	0	0%
ALL CAPS	0	0%
TOTAL	83	100%

CORE SUBSET

CAP STYLE	N	%
SENTENCE CASE	35	83%
ELEMENTS ONLY CAP	5	12%
MIXED	2	5%
OTHER	0	0%
ALL CAPS	0	0%
TOTAL	42	100%

Discussion and Recommendations: Attachments to this appendix, prepared to support Sections 1.5 Examples of High Quality Records and the discussion of P3-Aggregation, represent some variations in GILS record appearance. Investigators believe that standardization of record display, including typeface and weight will contribute to users' "footing" within GILS vs. other information space—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. It is recommended that decisions concerning standardization be based on published research concerning visual cues in human-machine interaction.

S4   Indentation

Findings: Roughly 1 in 4 records featured all flush-left text —i.e., no indentation was used to represent the element/subelement hierarchy.

TOTAL SAMPLE

INDENTATION	N	%
YES	61	73%
NO	22	27%
TOTAL	83	100%

CORE SUBSET

INDENTATION	N	%
YES	29	69%
NO	13	31%
TOTAL	42	100%

Discussion and Recommendations: The lack of indentation, when coupled with other style characteristics, such as no boldface, all capitals, etc., impedes scanning of record content for relevant terms. Investigators believe that standardization of record display, including indentation, will contribute to users' "footing" within GILS vs. other information space—e.g., recognition that GILS is a bounded (by function), top-down, two-dimensional service that spans across all agencies. It is recommended that decisions concerning standardization be based on published research concerning visual cues in human-machine interaction.

S5   Element Display Order

Findings: Roughly two-thirds of the total sample and nearly one-half of the Core subset records displayed elements in the order recommended by *FIPS Pub. 192* and the NARA *Guidelines*.

TOTAL SAMPLE

PREFERRED DISPLAY ORDER	N	%
YES	53	64%
NO	30	36%
TOTAL	83	100%

CORE SUBSET

PREFERRED DISPLAY ORDER	N	%
YES	24	57%
NO	18	43%
TOTAL	42	100%

Discussion and Recommendations: Lack of consistent and predictable ordering of metadata elements is disconcerting to users and may inhibit recognition of relevant terms. It is recommended that record designers standardize and use an ordering scheme based on systematic analysis of various user communities' preferences and consider optional (on-the-fly) re-ordering or suppression of elements upon client command as is provided by Z39.50-compliant servers. See also C1-Number of Elements Per Record.

S6   Definition of Acronyms

Findings: This analysis considered incidence of acronyms anywhere in the record, including the Acronym subelement. Only 12% of records containing acronyms failed to define them.

TOTAL SAMPLE

ACRONYMS	N	%
DEFINED	19	23%
NOT DEFINED	10	12%
NOT USED	54	65%
TOTAL	83	100%

CORE SUBSET

ACRONYMS	N	%
DEFINED	13	31%
NOT DEFINED	5	12%
NOT USED	24	57%
TOTAL	42	100%

Discussion and Recommendation: The incorporation of defined acronyms in government information undoubtedly assists users in NIDR. The absence of acronyms altogether in more than half of the records sampled was a surprising result, and record creators should be trained not to sacrifice relevant acronyms for record brevity. Section C3.1-Titles addresses the use of AIS acronyms.

S7   Citation of Legislation
Findings: Roughly half of the records sampled contained a reference to legislation concerning the information resource and/or its provision, including one instance of reference to GILS in Agency Program.

TOTAL SAMPLE

LEGISLATIVE CITE	N	%
YES	40	48%
NO	43	52%
TOTAL	83	100%

CORE SUBSET

LEGISLATIVE CITE	N	%
YES	21	50%
NO	21	50%
TOTAL	42	100%

Discussion and Recommendation: Inclusion of references to pertinent legislation improves NIDR only when the user searches in the "correct" (as used by the record creator) format—e.g., "Pub. L. 103-40" vs. "public law 103"). Given the fairly high incidence of legislative citation, it is recommended that a qualitative analysis of the citation formats be performed to determine the extent of variations. If warranted, a standard format should be incorporated into the NARA *Guidelines*. The feasibility of automating the insertion of substantively correct and properly formatted legislative citations upon inclusion of a program or project name (i.e., installation of an expert-system legislation index cross-reference macro) could reduce record creation and maintenance burden as well as facilitate identification of legislation-dependent resources for agency users (IRM and public information office personnel), Congressional users, and public policy researchers.

S8   Locally Defined Elements
Findings: This analysis considered locally defined elements as any not specifically mentioned in the NARA *Guidelines*. The 12 identified (labeled) locally-defined elements are listed below. (Note: some values were null.)

- URL
- System Products Disposition
- Organization [appears between Title and Abstract in addition to and of equal weight with Originator]
- Material Type
- Creation Date
- Disposition
- [Agency] Storage Authorized
- Disposition Authority
- Discription [sic]
- Record-Type
- Status
- Date Of Last Review

Discussion and Recommendations: The NARA *Guidelines* provide for use of locally defined elements in organizational elements and "when agencies wish to convey to the public or use for internal purposes information that is not part of the GILS Core. Whenever possible, such supplemental information should be associated with one of the GILS Core Elements as a locally-defined subelement to the Core Element [rather than in Supplemental Information]." It is recommended that the incidence and values of locally-defined elements be studied systematically; high incidence may indicate the need for an additional "official" element and extensive usage of a locally defined element that definitionally belongs with an extant element may dictate a clarification of the NARA *Guidelines* usage instructions.

5.0.   SUMMARY OF RECOMMENDATIONS

The following recommendations have been extracted from Section 4.0 Findings, Discussion, and Recommendations and arranged according to whether they are best implemented (1) at the agency level or (2) standards level by means of revision to the NARA *Guidelines*.

5.1.   Agency Level

1. Devise a hard-/software independent template and/or HTML editor for record formatting, or limit formatting responsibility to agency or subcontracted personnel with Web browsers.
2. Use machine-based spell checkers, or assign checking responsibility to someone other than the writer.
3. Use "auto-correct/auto-format" macros to standardize dates.
4. Audit the accuracy of Distributor and Points of Contact element values.
5. Enlist an objective, third-party (i.e., other than resource creator and/or record creator) to evaluate Title descriptiveness.
6. Automate the insertion of substantively correct and properly formatted legislative citations upon inclusion of a program or project name (i.e., install an expert-system legislation index cross-reference macro).
7. Develop training and complementary procedures for record creators to recognize the need for and to obtain the information for populating the Methodology, Sources of Data, and Schedule Number elements.
8. Use computer-generated time/stamp dates as a final "record checked and approved for release" procedure to promote responsibility for record content quality, facilitate audit sampling, and ease record maintenance.
9. Provide record creators and quality checkers a short, straightforward, procedural set of record-creation instructions as well as a FAQ list, pocket guide reference, context-sensitive online help, etc. to select elements required to describe the resource and its availability appropriately.
10. Monitor the incidence and values of new or "truly" locally-defined as well as those more correctly used as an extant element subelement to inform revision of the NARA *Guidelines*.
11. Implement Z39.50-compliant servers and clients, which will present customized views of the record through re-ordering or suppression of elements upon client command (e.g., allowing the user to select presentation of the Abstract with the Title in the results list to assist in relevance judgment).
12. On GPO Access GILS, provide a hypertext link from "*US Government Manual*" in the field definition files to a recast version of *Government Manual* Appendix A: Commonly Used Abbreviations and Acronyms
13. Participate in and pursue PURL and similar technology development efforts to ensure satisfaction of GILS unique requirements

5.2.   Standards Level (Revision of the NARA *Guidelines*)

1. Content analysis of the *Guidelines* by a third-party (i.e., not someone who was involved with the writing) is recommended to reveal areas of redundancy and ambiguity—e.g.,
    a. Restate the *Guidelines* Distributor element requirement from "Complete as many of the subordinate fields as necessary to identify the party" to "Complete all subordinate fields as available to assist the user in communicating with the Distributor."
    b. More clearly differentiate between placement of certain constraints (e.g., "requires forms-capable browser" in Access Constraint vs. Order Process)
    c. More clearly differentiate expected values between Purpose and Agency Program
    d. Clarify the rationale for isolating mandatory use of Sources of Data to describe AISs
2. Rename the Date of Last Modification element "Record Revision Date" and eliminate confusing "entry" nomenclature in the usage guidelines
3. Revise the definition, usage, and structure of the Resource Description subelement to provide physical description of the object as recognizable by the *user* rather than by the distributor; move this information

from the Abstract element to the Resource Description subelement; and make the (redefined) Resource Description subelement mandatory and structurally associated with the Abstract element.

4. Provide the additional GILS element "Object Represented" in order to evaluate aggregation, or "informational distance" from satisfaction of the information requirement and increase retrieval precision.

5. Standardize record display, including type font, weight, and size, as well as indentation and capitalization to "moor" users in GILS information space and promote the concept of a government-wide rather than agency-centric program. Base decisions on recognized research concerning visual cues in human-machine interaction.

6. Standardize and use an element display order scheme based on systematic analysis of various user communities' preferences.

7. Re-examine the concept and functionality of the GILS "Core." Should the requirement remain viable, clarify the rationale and guidance.

It is recommended that agencies and inter-agency oversight bodies implement the above recommendations with specific procedures, schedules, and performance measurements.

6.0.   FURTHER RESEARCH

Investigators discovered from this content analysis experience that the method might find optimal utility when employed in circumstances where specific user-defined criteria are known. For example, the user assessment technique described in Appendix C-5 Scripted Online User Assessment Methodology, for example, revealed that excessive record length (i.e., number of elements present) was off-putting to some users. The effects of record length on users for known-item searching vs. browsing, for example, could be studied by means of a record sample comprising a mix of popular resources (determined by Public Information Officers or transaction log analysis), newly created or newly aggregated resources, and resources sought frequently by professional intermediaries (such as federal depository librarians) but not end users. An understanding of how users read, evaluate, and "use" GILS records could inform the creation of customized record views.

The following list presents areas for further research of record content as an indicator of how well GILS is meeting expectations of users.

1. Hypertext incidence and placement (see Appendix C-5 and Appendix D-5 Online Scripted User Assessment Methodology and Instrument, respectively)

2. Government-wide frequency of change and clustering of Dates of Last Modification

3. Effects of "familiar"(e.g., LCSH) and specialized Controlled Vocabularies on *NIDR vis-a-vis* users' dependency on knowledge of agency mission and to increase precision of information retrieval (this effort should be in concert with the Library of Congress and GPO's Cataloging Branch)

4. User preferences for "catalog- or browsing based" NIDR (e.g., the approach of Yahoo! and the Argus Clearinghouse)

5. Appropriateness of content and placement of element values; a possible research project might isolate a random sample of selected data values and test whether GILS-cognizant vs. noncognizant users can place them in the "correct" (per published standards) elements

6. Record and resource aggregation effects on NIDR in terms of distance from satisfaction of an information need.

7. Control Identifier values as representing resource accession, IRM, or subject/object classification schemes.

8. Content of Web pages linking *to* GILS' homepages as a means to improve Local Subject Index Terms and Cross References through increased understanding of user expectations concerning the scope or nature of the GILS record collection. For example, links predominately via Web sites of the legal community may permit inclusion of more specific legal terms as well as citation of applicable cross-references. One current method of acquiring linking-from data is to execute an AltaVista <http://www.altavista.digital.com> search along the lines of link:http://www.[agency name].gov/gils - host:http://www.[agency name].gov. (Instructions are provided at the Alta Vista site.).

9. Evolution of data elements over time (as documented in working group minutes, listservs, standards, user-based research results, and agency-level training materials and procedures) as elucidating drivers for change—e.g., international, Federal, and state information policy; technology; standards; economics; nature of the resources; information life cycles; user expectations; agency mission; etc.

10. Content analysis of maintenance-intensive metadata (e.g., organizational subelements, Availability element URLs, cost, etc.) to reveal means of consolidating or otherwise arranging such data for ease or possible automation of update or record archiving.

11. Incidence of (1) a single information resource being "claimed" by multiple Record Sources—e.g., a CD-ROM content creator, content compiler, manufacturer, distributor and (2) "nonoriginal" Control Identifiers as an indicator of need for consensual policies on record-creation responsibility to avoid duplication of efforts and unnecessary maintenance burden.

These areas of research are considered viable in terms of improving GILS synergy with other products and services in government information space and its efficacy in NIDR, as well as in developing technical and procedural standards to guide the evolution of GILS.

7.0. CONCLUSION

The content analysis of GILS records provided valuable data concerning their accuracy, completeness, and serviceability in NIDR. In addition, it afforded a high-level view of the types and aggregation levels of resources being described as well as modes of packaging and dissemination.

Placed in context with the online user assessment results (see Appendix E-3 Scripted Online User Assessment Findings, Discussion, and Recommendations), the assessed degree of variation in the quality of GILS records demands immediate attention at both the agency and standards levels. Agencies can adapt the content analysis methodology to appraise the quality of their current GILS records as well as to identify early and act on evolving issues—e.g., such as scope of collection and levels of granularity—at the collection level. In addition, when applied to a government-wide sample, adaptations of this record content analysis will inform oversight bodies about levels of adherence to standards and synergy of GILS with other Federal government NIDR mechanisms.

Attachments E2-1a through E2-1d
Examples of High-Quality Records From the Sample

Four records exhibiting characteristics of "high quality" as defined by the record content analysis are reproduces in Attachments 1a through 1d. They have been reproduced as closely as possible to depict their actual display to an online user of GILS (i.e., font attributes, line spacing and indentation, etc.):

- *AHCPR Publications Clearinghouse* available at <http://www.dhhs.gov/progorg/oirm/newhhsgils.htm> by searching Control Identifier (quotes required) "HHS-AHC-00509"

- *Aviation Accident Synopses World Wide Web Page* available by browsing <http://www.ntsb.gov/Info/Info.htm> or directly at <http://www.ntsb.gov/Info/GILS/GILSSYN.htm>

- *Farm Credit Administration's Privacy Act Systems* available <http://www.access.gpo.gov/su_docs/gils/gilsfld.html> by searching Control Identifier (quotes required) "FCA/PA-1"

- *FEMA Publications Catalog* available <http://www.access.gpo.gov/su_docs/gils/gils.html> by searching Federal Emergency Management Agency for "FEMA0001"

These records contain mandatory elements populated with NARA *Guidelines*-compliant values and are highly readable and descriptive without excessive length. In addition, they represent a range of "information objects" and "containers": an information resource organization, an aggregated set of reports available via Web site, a "system of records" available via GPO Access, and a traditional printed publications catalog, respectively.

Title: AHCPR Publications Clearinghouse

Acronym: AHCPRPC

Originator: Department of Health and Human Services (HHS)
                U.S. Public Health Service (PHS)
                Agency for Health Care Policy and Research (AHCPR)

Local Subject Index: US Federal GILS, Clearinghouse

Abstract: The AHCPR Publications Clearinghouse is the primary
          storage and distribution point for all AHCPR
          publications. The Clearinghouse also maintains and
          manages AHCPR's automated mailing/inventory control
          system and manages the storage and shipping of AHCPR
          exhibits.

Purpose: The AHCPR Publications Clearinghouse serves as the
          Agency's direct mail and fulfillment service point.

Agency Program: A component of the Center for Health
                    Information Dissemination, this clearinghouse
                    is authorized under Section 6013, P.L. 101-
                    239, Omnibus Reconciliation Act of 1989.

Time Period of Content:
     Time Period Structured: 19940901 -
     Time Period Textual: 1 September 1994 - ongoing

Availability:
  Distributor:
     Name: Erin Henderson, Project Director
     Organization: Logistics Applications, Inc.
     Street Address: 9475 Gerwig Lane, Suite V
     City: Columbia
     State: MD
     Zip Code: 21046-1506
     Country: USA
     Network Address: None
     Hours of Service: 9 a.m.- 5 p.m., Monday - Friday
                        closed Federal Holidays
     Telephone: 800-358-9295; 301-621-3033 (local and non-U.S.)
     Fax: 410-290-3841
     TDD: 888-586-6340

     Order Process: AHCPRPC can be accessed via telephone, fax,
     TDD, or the Internet via the AHCPR Web Home Page.

Technical Prerequisites: For WWW access, Internet access and a
                         browser is required.

  Available Time Period:
     Electronic information access: 1 September 1994 - ongoing

```
   Available Linkage:
       Connect to AHCPRPC via AHCPR Home Page:
            http://www.ahcpr.gov/

   Available Linkage Type: plain text

Access Constraints: None

Use Constraints: None

Point of Contact:
       Name: Steven Merrill, Federal Project Officer;
             Judy Wilcox, Federal Alternate Project Officer
       Organization: Agency for Health Care Policy and Research
       Street Address: 2101 E. Jefferson Street, Suite 501
       City: Rockville
       State: MD
       Zip Code: 20852
       Country:  USA
       Network Address: smerrill@po5.ahcpr.gov;
                        awilcox@po5.ahcpr.gov or
                        puborder@po5.ahcpr.gov
       Hours of Service: Monday - Friday, 7 a.m.- 4 p.m.
       Telephone: Steven Merrill:  301-594-1364, extension 1350;
                  Judy Wilcox:     301-594-1364, extension 1389
       Fax: 301-594-2286

Control Identifier: HHS-AHC-00509

Record Source: Agency for Health Care Policy and Research
               Center for Health Information Dissemination

Date of Last Modification: 19960708
```

<div align="right">

Attachment E2-1b
High Quality Record From the Sample

</div>

Title: Aviation Accident Synopses World Wide Web Page

Originator:

Department/Agency Name: National Transportation Safety Board

Major Organizational Subdivision: Office of Research and Engineering

Name of Unit: Analysis and Data Division

Local Subject Index:

Local Subject Term: US Federal GILS

Abstract: The Aviation Accident Synopses World Wide Web Page contains short reports describing aircraft accidents and incidents and their probable cause, and contributing factors. Included are civil aviation accidents within the United States, its territories and possessions, and in international waters. Incidents (mishaps that do not meet the aircraft damage or personal injury thresholds in the regulatory definition of "accident") investigated by the National Transportation Safety Board are contained in the database in the same form as accidents. Reports are available for the time period 1983 to the present. Generally, a preliminary report is available on line within a few days of an accident. When the investigation is completed, the preliminary report is replaced with a final description of the accident and its probable cause. The World Wide Web page provides access to more than 35,000 reports through 1995, and is growing by approximately 2,250 cases per year. Access to specific accidents and incidents is by means of monthly lists of all such occurrences in the National Transportation Safety Board's Aviation Accident Data Base. Synopses and monthly lists are updated daily.

Purpose: The Aviation Accident Synopses World Wide Web page provides to the public direct access to limited information regarding aviation accidents investigated by the National Transportation Safety Board.

Agency Program: The National Transportation Safety Board provides information to the public via its World Wide Web page with the intent of making safety related information rapidly and easily available to its customers worldwide.

Spatial Reference:

Geographic Name:

Geographic Keyword Name: United States

Geographic Keyword Name: United States Territories and Possessions

Geographic Keyword Name: International Waters

Time Period of Content:

Time Period-Structured: 1983-

Time Period-Textual: 1983 - [ongoing]

Availability:

Distributor:

Name: NTSB WWW Server

Organization: National Transportation Safety Board

Network Address: webmaster@ntsb.gov

Order Process: Accident synopses are available via the internet using an http client. From the NTSB Home Page
(www.ntsb.gov), select "Aviation" then "Accidents".

Technical Prerequisites: Access to the Internet and an http client.

Available Linkage: http://www.ntsb.gov

Available Linkage Type: text/plain

Sources of Data: Synopses are produced from data developed by aviation accident investigators of the National Transportation Safety Board.

Access Constraints: Synopses may be located only by searching monthly lists of accidents and incidents.

Use Constraints: Public Law 93-633 and the National Transportation Safety Board's regulations 49 CFR 835 prohibit the use of accident/incident findings, including the probable cause and contributing factors as evidence in any suit or action for damages arising from that event.

Point of Contact:

Name: Analysis and Data Division (RE-50)

Organization: National Transportation Safety Board

Street Address: 490 L'Enfant Plaza East

City: Washington

State: DC

Zip Code: 20594-2000

Country: USA

Network Address: webmaster@ntsb.gov

Hours of Service: 8:00 a.m. - 5:00 p.m.

Telephone: 202-314-6550

Telephone: 800-877-6799

Fax: 202-314-6598

Cross Reference:

Cross Reference Title: Aviation Accident Database

Cross Reference Linkage: GILS: NTSB0001

Cross Reference Type: text/html

Cross Reference Title: Safety Recommendations Database

Cross Reference Linkage: GILS: NTSB0002

Cross Reference Type: text/html

Cross Reference Title: Accident Investigation Dockets

Cross Reference Linkage: GILS: NTSB0003

Cross Reference Type: text/html

Cross Reference Title: Formal Reports

Cross Reference Linkage: GILS: NTSB0004

Cross Reference Type: text/html

Schedule Number: Not scheduled

Control Identifier: NTSB0005

Record Source:

Department/Agency Name: National Transportation Safety Board

Major Organizational Subdivision: Office of Research and Engineering

Name of Unit: Analysis and Data Division

Date of Last Modification: 19960329

TITLE: Farm Credit Administration's Privacy Act Systems

Originator: Farm Credit Administration

LOCAL SUBJECT INDEX: US Federal GILS

ABSTRACT: In compliance with the Privacy Act of 1974, the Farm Credit Administration publishes notices in the Federal Register about the record systems the agency maintains that are retrieved by name or personal identifier. These record systems are commonly referred to as "Privacy Act systems" and the information published about them are referred to as "system notices". The Privacy Act systems maintained by FCA primarily cover FCA employees. Each "system notice" contains the following information: system name; system location; categories of individuals covered by the system; categories of records in the system; authority for maintenance for the system; routine use of records maintained in the system; policies and practices for storing, retrieving, accessing, retaining, and disposing of records in the system; the name and address of the system manager; and procedures for access and contesting information in the records. After publication in the Federal Register, FCA's system notices are compiled in Privacy Act Issuances, a biennial compilation of all Federal agency Privacy Act notices, which has been published by the Government Printing Office in CD-ROM format since 1993. A free public-access version of Privacy Act Issuances is available on the Internet. Any revisions to FCA's Privacy Act notices since the last compilation can be found by reviewing the Federal Register.

PURPOSE: FCA's Privacy Act system notices provide the public with information about systems of records maintained by the agency that are retrieved by name or personal identifier.

AGENCY PROGRAM: 5 USC 552a, the Privacy Act of 1974

AVAILABILITY:

Distributor:

Name: Superintendent of Documents, U.S. Government Printing Office (GPO)

Street address: Mail Stop SDE, 732 North Capitol Street NW

City: Washington, D.C

Zip Code: 20401

Country: USA

Telephone: 202-512-1530

Fax: 202-512-1262

Hours: 7:00 a.m. - 5:00 p.m. M-F EST

Resource: Privacy Act Issuances (on-line version)

Order Process: Available through the Internet free of charge or through dial-up access. There is no charge for using the phone line, but you may incur long distance charges.

Technical Prerequisites: Access to Internet and an http browser or telnet or WAIS client or access to a modem and telephone line.

Linkage: http://www.access.gpo.gov/sudocs/gils/gils.html; telnet: swais.access.gpo.gov; wais: wais.access.gpo.gov; asynchronous: 202- 512-1661

Linkage Type: text/plain

ACCESS CONSTRAINTS: None.

USE CONSTRAINTS: None.

POINT OF CONTACT:

Name: Office of General Counsel

Organization: Farm Credit Administration

Street address: 1501 Farm Credit Drive

City: McLean

State: Virginia

Zip code: 22102-5090

Country: USA

Telephone: 703-883-4022

CONTROL IDENTIFIER: FCA/PA-1

RECORD SOURCE: Farm Credit Administration

DATE OF LAST MODIFICATION: 19970220

Attachment E2-1d
High Quality Record From the Sample

```
Title:  FEMA Publications Catalog

Originator:
     Department/Agency Name:  Federal Emergency Management Agency

Abstract:  This catalog provides information regarding publications and
other printed matter produced by FEMA.

Purpose:  This catalog provides directions/addresses of where and how to
obtain copies of FEMA publications and other printed matter produced by
FEMA.

Agency Program:  These publications are made available as a customer
service to the public.  The publications also document FEMA's mission and
programs that are available to the public, such as emergency management
training.

Availability:
     Distributor:
          Distributor Name:  Federal Emergency Management Agency
          Distributor Organization:  FEMA Publications Warehouse
          Distributor Street Address:  P. O. Box 2012
          Distributor City:  Jessup
          Distributor State:  MD
          Distributor Zip Code:  20794-2012
          Distributor County:  USA
          Distributor Network Address:  None
          Distributor Hours of Service:  8:30 a.m. to 5:00 p.m. EST
          Distributor Telephone:  1-800-480-2520
          Distributor Fax:  301-497-6378

Order Process:  Requests for copies of this catalog can be made by writing
or calling and asking for FEMA Publication 20.  Requests are limited to 100
copies.  Any of the publications listed in the catalog can be ordered by
following the instructions listed in the catalog.

Access Constraints:  There are no access constraints for the catalog for
requests from the United States; however, individual publications listed in
the catalog may indicate some access constraints.  Foreign requests must be
approved by the Office of Security, FEMA, prior to being fulfilled.

Use Constraints:  None.

Point of Contact:
     Contact Name:  Printing & Publications Branch
     Contact Organization:  Program Services Division, FEMA
     Contact Street Address:  500 C Street, SW
     Contact City:  Washington
     Contact State:  DC
     Contact Zip Code:  20472
     Contact Country:  USA
     Contact Network Address:  None
     Contact Hours of Service:  8:30 a.m. to 5:00 p.m. EST
     Contact Telephone:  202-646-2650
     Contact Fax:  202-646-3524

Schedule Number:  Scheduled:  N1-311-86-1 1K6

Control Identifier:  FEMA0001
```

Record Source: Federal Emergency Management Agency

Date of Last Modification: 19950404

Attachments E2-2a through E2-2d
Aggregation Code Examples

These attachments illustrate the operational definitions of record and resource aggregation used in the record content analysis; no other evaluative criteria apply.

The following four aggregation codes, discussed more fully in Appendix C-4 Record Content Analysis Methodology, were applied to all sampled records:

Code	Operational Definition	Examples
Record Aggregates Objects	GILS record, by virtue of its creation, collects discrete information resources that record content indicates would not have otherwise been collected or aggregated. Assigned in the absence of clues within the record that the represented objects were heretofore packaged *as this collection* to optimize information discovery and retrieval.	• Privacy Act Systems compilation • files • press releases • forms
Aggregated Object Represented	GILS record represents an *a priori* or purposeful collection of information resources—e.g., woodpecker database or agency website. GILS record represents an object that collects, or comprises, two or more discrete information objects, and that represents a collection of standalone information files or products packaged together on the basis of a common theme or subject for functional convenience.	• CD-ROM of regulations • System that compiles Privacy Act records • job line of open requisitions
Discrete Object Represented	GILS record describes a standalone document-level entity that does not meet the criteria for "object aggregates metadata" below.	• annual report • videotape
Object Aggregates Metadata	GILS record describes a pre-existing metadata collection, or "locator," as an information resource.	• directory • catalog • index • log

Title: Briefing Materials, Public Comments, other Related
     Official Files
     Acronym:

Originator:
     Department/Agency Name: Consumer Product Safety Commission

Local Subject Index:
     Local Subject Term: US Federal GILS
     Local Subject Term: Consumer Product Safety

Abstract: These files include the briefing memoranda prepared by
the staff and reviewed by the Commissioners and public comments
to any rulemaking documents.

Purpose: To adhere to the provisions of the Consumer Product
Safety Act and other laws administered by the Commission.

Agency Program: To store, index and maintain the records
associated with official Commission actions.

Availability:
     Distributor Name: Office of the Secretary, Freedom of
     Information Division
     Organization: Consumer Product Safety Commission
     Street Address: 4330 East West Highway
     City: Bethesda
     State: MD
     Country: USA
     Zip Code: 20207
     Telephone: 301-504-0800
     Fax: 301-504-0127
     Order Process: Official records are available in hard copy
     from the Office of the Secretary or the Office of
     Information and Public Affairs at the address provided..

Availability:
     Distributor Name: Office of Information and Public Affairs
     Organization: Consumer Product Safety Commission
     Street Address: 4330 East West Highway
     City: Bethesda
     State: MD
     Country: USA
     Zip Code: 20207
     Telephone: 301-504-0785
     Fax: 301-504-0862
     Order Process: Official records are available in hard copy
     from the Office of the Secretary or the Office of
     Information and Public Affairs at the address provided.

Access Constraints: None

Use Constraints: None

```
Point of Contact:
     Name:  Office of the Secretary
     Organization:  Consumer Product Safety Commission
     Street Address:  4330 East West Highway
     City:  Bethesda
     State:  MD
     Zip Code:  20207
     Country:  USA
     Network Address:  info@cpsc.gov
     Hours of Service:  8:30 a.m. - 5:00 p.m.
     Telephone:  301-504-0800
     Fax:  301-504-0127

Control Identifier:   CPSC0014

Record Source:
     Department/Agency Name:  Consumer Product Safety Commission

Date of Last Modification:  19960510
```

TITLE: Worldwide Real Property Inventory System

ACRONYM: WRPIS

ORIGINATOR

DEPARTMENT/AGENCY NAME: General Services Administration (GSA)

MAJOR ORGANIZATIONAL SUBDIVISION: Public Building Service (PBS)

MINOR ORGANIZATIONAL SUBDIVISION: Office of the Chief Information Officer

NAME OF UNIT: Customer Service Division

ABSTRACT: The Worldwide Real Property Inventory System (WRPIS) supports the worldwide inventory information reporting cycle which involves collecting real property data (GSA FORM 1166) from Executive Branch Federal agencies, analyzing the data, and providing worldwide real property data to the public, private organizations, Congress, and other Federal agencies by providing a single source of information for both owned and leased property. The WRPIS system accepts data from Foundation Information for Real Property Management System (FIRM) as well as other media from the Executive Branch agencies. Reports and information available from WRPIS include the Summary Report of Real Property Leased by the United States Throughout the World, the Summary Report of Real Property Owned by the United States Throughout the World, and the World Wide Geographic Location Code Table.

BEGIN DATE: 1993

END DATE:

AVAILABILITY

DISTRIBUTOR

DISTRIBUTOR NAME: Public Buildings Service (PBS)

DISTRIBUTOR ORGANIZATION: General Services Administration (GSA)

DISTRIBUTOR STREET ADDRESS: 18th and F Streets, NW

DISTRIBUTOR CITY: Washington

DISTRIBUTOR STATE: DC

DISTRIBUTOR ZIP CODE: 20405

DISTRIBUTOR COUNTRY: USA

DISTRIBUTOR NETWORK ADDRESS:

DISTRIBUTOR HOURS OF SERVICE: 8:00 a.m. - 4:30 p.m.

DISTRIBUTOR TELEPHONE: 202-501-0856

DISTRIBUTOR FAX:

RESOURCE DESCRIPTION:

ORDER PROCESS:

ORDER INFORMATION: There is no on-line access to WRPIS outside of the Public Building Service except for the Worldwide Geographic Location Code Table. This information is accessible through the GSA Electronic Management Information (GEMI) Bulletin Board, 202-219-0132. This data is also available by printout and/or electronic media. To request printouts or electronic copies of the Worldwide Geographic Location Code Table or copies of the annual publications "Summary Report of Real Property Leased by the United States Throughout the World" and "Summary Report of Real Property Owned by the United States Throughout the World" write to the PBS Office of Governmentwide Real Property Policy (PG) at the address shown in the Distributor field.

COST:

COST INFORMATION: Some reports are free to other Federal agencies but there may be charges for the public and private sector. Specific costs will be determined on a case by case basis based on the specific requirements of the request.

TECHNICAL PREREQUISITES: To access GEMI Bulletin Board, 2,400 to 14,400 BAUD modem, ANSI or VT1000 terminal emulation, communications software

AVAILABLE TIME PERIOD

TIME PERIOD STRUCTURED:

TIME PERIOD TEXTUAL:

AVAILABLE LINKAGE: To access GEMI Bulletin Board, dial 202-219-0312, settings N-8-1-F.

AVAILABLE LINKAGE TYPE:

SOURCES OF DATA: Executive Branch agencies

ACCESS CONSTRAINTS

GENERAL ACCESS CONSTRAINTS: None.

DOCUMENTATION:

ORIGINATOR DISSEMINATOR CONTROL:

SECURITY CLASSIFICATION CONTROL:

USE CONSTRAINTS: None

POINT OF CONTACT FOR FURTHER INFORMATION

CONTACT NAME: Office of the Chief Information Officer

CONTACT ORGANIZATION: Public Building Service (PBS)

CONTACT STREET ADDRESS: 18th and F Streets, NW

CONTACT CITY: Washington

CONTACT STATE: DC

CONTACT ZIP CODE: 20405

CONTACT COUNTRY: USA

CONTACT NETWORK ADDRESS:

CONTACT HOURS OF SERVICE: 8:00 a.m. - 4:30 p.m.

CONTACT TELEPHONE: 202-501-9170

CONTACT FAX: 202-208-7087

PURPOSE: See abstract

AGENCY PROGRAM: WRPIS supports the Federal Management Regulations 41 CFR 101-3 provision which requires Executive Branch Agencies, on an annual basis, to submit information regarding their properties to GSA.

SCHEDULE NUMBER: Not scheduled.

CONTROL IDENTIFIER: GSACP9002

RECORD SOURCE

RECORD SOURCE/DEPARTMENT AGENCY NAME: General Services Administration (GSA)

RECORD SOURCE/MAJOR ORGANIZATIONAL SUBDIVISION: Information Technology Service (ITS)

RECORD SOURCE/MINOR ORGANIZATIONAL SUBDIVISION: Office of GSA-Wide Information Technology (IT)

RECORD SOURCE/NAME OF UNIT: Center for GSA-Wide IT Systems Planning and Management

DATE OF LAST MODIFICATION: 19951031

SUPPLEMENTAL INFORMATION:

SYSTEM PRODUCTS DISPOSITION:

LOCAL SUBJECT INDEX

LOCAL SUBJECT TERM: US Federal GILS

Title: Investment Funds Brochure

Originator: U.S. International Development Cooperation Agency
(IDCA)/Overseas Private Investment Corporation (OPIC)

Abstract: The Investment Funds Brochure contains general information
on each of the Investment Funds guarantied by OPIC. For each fund,
the Fund Manager, Capital, Status of Funds Availability, Target
Market, Targeted Sectors, Preferred Investment, Exit Strategy,
Requirements, Qualifications for U.S. Business Participation, and
Additional Services are listed.

Purpose: The Investment Funds Brochure is designed to provide a brief
synopsis of OPIC's guarantied funds to help American business
executives and entrepreneurs interested in investing.

Agency Program: The Investment Funds Brochure describes Guarantied
Funds managed by the Investment Funds Program.

Availability:
    Distributor Name: Information Officer
   Organization: Overseas Private Investment Corporation
   Street Address: 1100 New York Avenue, N.W.
    City: Washington
        State: D.C.
     Country: U.S.A.
 Zip Code: 20527-0001
    Network Address: opic/s=info@mhs.attmail.com
    Hours of Service: 8:45 a.m. to 5:30 p.m.
        Telephone: 202/336-8799
 Fax: 202/336-8700
      Order Process: The Investment Funds Brochure is available
without charge by writing to the Information Officer, at the above
address, or by faxing a request to the number above, or by calling the
Information Line at the phone number above.

Availability:
      Distributor Name: Depository Library Program
   Organization: Government Printing Office
       Order Process: Item number OP.1.2; depository item number
     0834-W-02.

Sources of Data: The Investment Funds Brochure provides information
from sources within OPIC.

Access Constraints: None.

Use Constraints: None.

Point of Contact:
     Distributor Name: Information Officer
   Organization: Overseas Private Investment Corporation
   Street Address: 1100 New York Avenue, N.W.
    City: Washington

```
        State: D.C.
      Country: U.S.A.
 Zip Code: 20527-0001
     Network Address: opic/s=info@mhs.attmail.com
     Hours of Service: 8:45 a.m. to 5:30 p.m.
         Telephone: 202/336-8799
 Fax: 202/336-8700
```

Schedule Number: Scheduled- N1-420-93-1, #26.

Control Identifier: IDCA/OPIC-GILS: 0008

Record Source:
   Agency Name: U.S. International Development Cooperation
                        Agency (IDCA)/Overseas Private Investment
Corporation
                (OPIC)
 Major Organizational Subdivision: Management Services Department
        Name of Unit: Information Center

Date of Last Modification: 19951218

Title: Office of the General Counsel Library Catalog

Originator: Pension Benefit Guaranty Corporation (PBGC)/Office of the General Counsel

Controlled Vocabulary: Yes

Abstract: The Office of the General Counsel Library Catalog describes the library's holdings. Approximately 1400 titles in the area of pensions and pension law, bankruptcy, administrative law, and Federal practice. The library is a Federal Government Depository and maintains a small collection of government legislation, regulatory, and other documents in print and electronic format.

Begin Date: 1990

Purpose: The Office of the General Counsel Library Catalog enables the user to locate library resources and materials. The catalog is used as an automated finding aid.

Agency Program: Library resources support the work of agency staff.

Distributor:
Name: Office of the General Counsel
Organization: Pension Benefit Guaranty Corporation
Street Address: Suite 340, 1200 K Street
City: Washington, D.C.
State: N/A
Zip: 20005-4026
Country: USA
Telephone: (202) 326-4004
FAX: (202) 326-4112

Order Process: Currently, there is no on-line access to the General Counsel Library Catalog. The catalog is available to users in the OGC library which is open to the public during business hours (below).

Available linkage: PBGC Home page URL: http://www.pbgc.gov
Available linkage type: html

Sources of Data: Inventory of library holdings collected internally from PBGC departments, and outside government agencies and sources.

Use Constraints: None

Point of Contact:
Lilian H. Fry, Librarian
Office of the General Counsel
Organization: Pension Benefit Guaranty Corporation
Street Address: Suite 340 1200 K Street
City: Washington, D.C.

```
State: N/A
Zip: 20005-4026
Country: USA
Hours: 9:00 a.m. - 4:00 p.m.
Telephone: (202) 326-4004
FAX: (202) 326-4112

Schedule Number:   General Records Schedule Number #20.9.

Control Identifier: PBGC0003

Date of Last Modification: 19961119
```